

CUE-GÜTE JENSEITS VON VALIDITÄT UND DISKRIMINATIONSRATE

**BEITRÄGE ZUR SUCHE VON PROBABILISTISCHER
INFORMATION IM ENTSCHEIDUNGSPROZESS**

Abhandlung
zur Erlangung der Doktorwürde
der Philosophischen Fakultät
der Universität Zürich

vorgelegt von
SERAPHINA ZURBRIGGEN
Betten / VS

Angenommen im Herbstsemester 2008 auf Antrag von
Prof. Dr. Damian Läge und
Prof. Dr. Wolfgang Marx

Zürich 2008

INHALTSVERZEICHNIS

ZUSAMMENFASSUNG /ABSTRACT	II
ÜBERBLICK: INFORMATIONSSUCHE VON PROBABILISTISCHER INFORMATION IM ENTSCHEIDUNGSPROZESS.....	1
DIE INFORMATIONSKONSISTENZ	
Kapitel 1: Wie zufrieden machen Informationen? Die Bewertung von Informationen in Bezug auf Urteilssicherheit und Urteils- zufriedenheit	21
LERNUMWELTEN UND CUE-AUSPRÄGUNGSRATE	
Kapitel 2: Wie lernen wir zu diskiminieren?	41
ZEITLICH DYNAMISCHE UMWELTEN UND ADAPTIVITÄT	
Kapitel 3: Entscheiden in zeitlich dynamischen Umwelten: Suchen und Stoppen mit unterschiedlich stabilen Cues.....	69
Kapitel 4: Die zeitliche Dynamik von Umwelten: Heuristisches Wahlverhalten am Beispiel eines Trinkwasser-Szenarios mit und ohne Zeitachse	87
Kapitel 5: Ressource Zeit statt Geld: Ein Vergleich zweier Szenarien zur Entscheidungsforschung.....	107
Kapitel 6: Adaptivität der Informationssuche in statischen und in dynamischen Umwelten	117
KRITISCHE BEMERKUNGEN.....	
Kapitel 7: Die Präsentation von Cue-Gütemassen im Information Board Setting: Validität, Diskriminationsrate und Stabilitätsrate	141
DANKSAGUNG.....	151
CURRICULUM VITAE.....	152

ZUSAMMENFASSUNG

Entscheidungen benötigen Informationen. Diese Maxime, welche das Verhalten von Menschen in ihrer Umwelt besser macht als der Zufall, dürfte in der Forschung unbestritten sein. Und genauso unbestritten dürfte es sein, dass Menschen in der Lage sein müssen, a) sich diese Informationen zu beschaffen und b) sie so auszuwerten, dass sie auf dieser Basis die entsprechenden angemessenen Entscheidungen treffen können. Letzteres ist der traditionelle Fokus der Entscheidungspsychologie, nämlich diejenigen kognitiven Verfahren ausfindig zu machen, mit denen vorliegende Informationen in eine Entscheidung transformiert werden. Die Entscheidungsforschung fokussiert traditionell auf den Moment der Entscheidung selbst, wobei nun seit einigen Jahren auch die aktive Informationssuche vor einer Entscheidung mitmodelliert wird: Urteilsheuristiken im theoretischen Rahmen der Adaptiven Toolbox von Gigerenzer et al. (1991) legen fixe Such- und Stoppregeln nahe, anhand derer Informationsquellen sequentiell gemäss derer Validität und Diskriminationsrate befragt werden. In dieser Arbeit werden diese sequentiellen Such- und Stoppregeln genauer beleuchtet: Einerseits finden wir eine über die festgelegte Stoppregel hinausgehend adaptive Stoppregel in Bezug auf das Anspruchsniveau an Urteilssicherheit und auf die Informationskonsistenz von mehreren Informationen. Zudem finden wir über die festgelegten Suchreihenfolgen (z.B. in absteigender Validitätsfolge) hinausgehend eine adaptive sequentielle Informationssuche in Abhängigkeit der Qualität der nicht diskriminierenden Information in Richtung der entsprechenden Ausprägungsrate. Und zuletzt erweitern wir das klassische Börsenszenario um eine Zeitdimension (anhand einer Stabilitätsrate für Alternative und Cues) und führen ein neues Experimentalsetting, das Trinkwasser-Szenario ein, welches das Verhalten vor mehreren Entscheidungen messbar und vergleichbar macht und die Ressource Zeit mitberücksichtigt. Da eine aktive, sequentielle Informationssuche nach nützlicher Information Zeit benötigt, wird diese Zeit zu einem wichtigen Kostenfaktor. Diese Beiträge erweitern das theoretische Feld der Entscheidungspsychologie, indem sie den Prozess vor einer Entscheidung um eine Zeitdimension spezifizieren und die entsprechend der Toolbox geforderte Adaptivität des Verhaltens genauer beleuchten.

ABSTRACT

In order to make a decision people rely on information. This maxim which renders human behaviour better than chance is undisputed in research. It is furthermore unquestioned that humans have to be able to a) acquire necessary information and b) evaluate this information in order to make a decision. Only recent research shifted the focus from the actual moment of decision making to the active search for probabilistic information before a decision: decision heuristics in the theoretical framework of the Adaptive Toolbox (Gigerenzer et al., 1991) rely on decision rules, but also on determined searching and stopping rules for cue-information defined by validity and discrimination rates. Here, we take a closer look at these sequential searching and stopping rules: First, we find an adaptive stopping rule which is dependent on the desired level of confidence and the consistency of several pieces of information. Second, we find an adaptive sequential information search which is dependant on the quality of non-discriminative information. Finally, instead of the classical stock market experimental setting used for decision research in the past, a drinking water scenario is explored. This setting introduces the resource time by offering a stability rate for decision alternatives and cues. This allows the analysis of how information is dealt with when searching for reliable cues requires time. The present contribution broadens the theoretical field of decision psychology by taking into account the resource time and by specifying the adaptive behaviour postulated by the Toolbox.

ÜBERBLICK: INFORMATIONSSUCHE VON PROBABILISTISCHER INFORMATION IM ENTSCHEIDUNGSPROZESS

Einleitung

Wir Menschen sind in unserer heutigen Informationsgesellschaft umgeben von zahlreichen Informationen, wertvollen und nutzlosen, widersprüchlichen und bestätigenden, neuen und alten. Immer schwieriger wird es, die Bedeutung einer Information in einer aktuellen Situation zu eruieren und mehrere Informationen zu einem Urteil zu integrieren, um sich entscheiden zu können. Dass sich der Mensch bei dieser Verarbeitung von Informationen und beim Entscheiden nicht rational verhält, wurde in unzähligen Studien immer wieder belegt. Zum Beispiel interpretieren Tversky und Kahnemann (Tversky & Kahneman, 1974) im Sinne eines ökonomischen Menschenbilds das Scheitern der Personen als Fehlverhalten. Mit Simons Argumentation einer begrenzten Rationalität (Simon, 1955) wurde der Mensch jedoch wieder als vernünftiges Wesen rehabilitiert. In dem Artikel von Simon 1955 wird der Begriff *bounded rationality* noch nicht verwendet, aber die gedanklichen Bausteine der neuen ökonomischen Theorie nahmen dort in der Argumentation, dass der Mensch den Limitationen eines biologischen Organismus unterliegt, ihren Anfang. So entstand in den 1950 und 1960 zum Zeitpunkt der kognitiven Wende dieses Bild vom informationsverarbeitenden Menschen. Mit der Entdeckung der Evolution als treibende Kraft für die kognitive Grundausstattung des Menschen ist dann in den 60er und 70er Jahren die Überzeugung entstanden, dass soziale und kognitive Fähigkeiten stets im Kontext ihrer stammesgeschichtlichen Entwicklung zu modellieren seien. Mit der Metapher eines Sackmessers postulierten Cosmides und Tooby (1994) in dieser evolutionspsychologischen Tradition eine menschliche Psyche, die nicht ein allgemeiner, jedes Phänomen aufklärender Problemlöser ist, sondern eher einem flexiblen Überlebenswerkzeug gleicht, das sich adaptiv entwickelt hat. Mit einer ähnlichen Metapher findet in der Kognitionspsychologie die Folgerung, Fähigkeiten als adaptive Anpassungen an eine bestimmte Umwelt zu sehen, ihren klarsten Ausdruck in der Konzeption der „Adaptiven Toolbox“ von Gigerenzer, Todd und der ABC-Research Group (1999).

Die Adaptive Toolbox

In dem Modell der „Adaptiven Toolbox“ wird von sequentiellen Prozessen bei der Verarbeitung von Informationen ausgegangen: „In the kind of inference task we are concerned with, cues have to be searched for, and the mind operates sequentially, step by step and cue by cue“ (Martignon & Hoffrage, 1999, S.137). Verknüpft mit dieser Annahme von sequentiellen Prozessen postulierten Gigerenzer, Hoffrage und Kleinbölting (1991) das Probabilistische Mentale Modell (PMM), welches basierend auf dem Linsenmodell von Brunswik (1952) entwickelt wurde. Brunswik war davon ausgegangen, dass Menschen Objekte, Ereignisse und Zustände in der Umwelt nie direkt erfahren können, sondern nur durch die Konstruktion von mentalen Repräsentationen erschliessen. Aufgrund dieser mentalen Repräsentationen fällt man schlussfolgernde Urteile über die Umwelt. Das bedeutet, dass der Mensch bildlich gesprochen durch die Linse Beobachtungen macht und indirekte Hinweisreize, sogenannte Cues sammelt, anhand welcher er dann probabilistische Schlussfolgerungen über die Umwelt zieht. Das Probabilistische Mentale Modell beschreibt nun in diesem Kontext, wie Personen ein

mentales Modell einer Situation bilden, das die relevanten Umweltaspekte einschliesslich der Unsicherheit abbildet. Die indirekten Hinweisreize (z.B. ob eine Stadt eine Bundesliga hat oder nicht gilt als Prädiktor für deren Einwohnerzahl) werden auch Cue-Informationen genannt und stehen in korrelativem Zusammenhang mit der Zielvariable (hier der Städtegrösse). Gigerenzer et al. nehmen an, dass die ökologische Validität einer Cue-Information im PMM als eine bedingte Häufigkeit, mit welcher eine richtige Vorhersage auf die Zielvariable gemacht wird, repräsentiert wird und die Validität somit die massgebliche Grösse der Informationsgüte ist. Die „Adaptive Toolbox“ beinhaltet viele verschiedene Urteilsheuristiken, die bekanntesten sind die Rekognitionsheuristik (Goldstein & Gigerenzer, 1999, 2002), *Take The Best*/*Take The Last* und *Minimalist* (Gigerenzer, Todd, & Group, 1999). Allen gemeinsam sind streng sequentielle und regelbasierte Verarbeitungsschritte, welche anhand der drei Bausteine Suchregel, Stopregel und Entscheidungsregel definiert werden.

Ursprünglich definierten Gigerenzer und Kollegen im Probabilistischen Mentalen Modell die Cue-Validitäten im Bezug auf probabilistische Informationen im Gedächtnis: „A PMM connects the specific structure of the task with a probability structure of a corresponding natural environment (stored in long-term memory)” (Gigerenzer et al., 1991, S.507) und “In Brunswik’s view, cue validities are learned by observing the frequencies of co-occurrences in an environment“ (S.510). Ausschliessliches Mass der Informationsgüte ist im Modell von Gigerenzer & Goldstein (1996) die Validität der Cues, also deren Vorhersagekraft im Falle des Diskriminierens (des Erhaltens einer Information).

Die bekannteste und experimentell meist untersuchte Heuristik *Take The Best* postuliert eine lexikografische Suchregel gemäss den Validitäten der Informationen, wobei diese Validitäten vorhanden bzw. gelernt sein müssen. Die Suchregel von *Take The Best* lautet folgendermassen: „Choose the cue with the highest validity that has not yet been tried for this choice task. Look up the cue values of the two objects” (Gigerenzer & Goldstein, 1999, S.81). Wobei die Autoren betonen, dass die Reihenfolge dieser Cue-Validitäten nicht immer der optimalen entsprechen muss, sondern auch von einer kleinen Referenzmenge ausgehend grob geschätzt werden kann. Nun bemerkten aber schon Martignon & Hoffrage (1999), dass ein Cue nicht nur nützlich ist, wenn er gute Vorhersagen macht, sondern auch wenn er häufig gebraucht werden kann, weil er zwischen zwei Alternativen diskriminiert, d.h. also eine hohe Diskriminationsrate hat. So postulierten sie als Gütemass der Cues eine Kombination von Validität und Diskriminationsrate: den Success. Sucht jemand die Informationen gemäss Success, dann wird davon ausgegangen, dass nach dem ersten Cue die Informationssuche gestoppt und bei nicht diskriminierender Information geraten wird.

Diese Theorie der adaptiven, streng sequentiellen Heuristiken wurde zahlreich anhand von Simulationsstudien überprüft (z.B. Czerlinski, Gigerenzer, & Goldstein, 1999; Hogarth & Karelaia, 2006) und deren Überlegenheit gegenüber vollständigen Informationsintegrationsmodellen belegt. Als klassisches Beispiel wurde dazu immer wieder das Städteparadigma verwendet (Gigerenzer & Goldstein, 1999). Beim Städteparadigma wird anhand von verschiedenen probabilistischen Cues, wie z.B. dem Vorhandensein einer Hauptstadt, einer Bundesliga oder einer Universität, auf die grössere von zwei Städten getippt. Doch überprüfbare Verhaltensdaten lassen sich, abgesehen von Zeitmessung von Gedächtnisabruf (z.B. Bergert & Nosofsky, 2007; Bröder & Gaissmaier, 2007) nur durch aufwändige empirische Lernexpe-

rimente mit im Experiment gelerntem Wissen über die Cue-Gütemasse erwerben. Letzteres Vorgehen wird immer häufiger angewendet, um dem Postulat von Gigerenzer (1999), es handle sich um Heuristiken in Bezug auf *inferences from memory* (versus *inferences from givens*) gerecht zu werden.

Weitere Modelle

Neben diesem Modell der „Adaptiven Toolbox“ entwickelten andere Autoren weitere Modelle von sequentieller Verarbeitung, die nicht von einem einfachen regelbasierten Verarbeiten von Informationen, sondern von *evidence accumulation* ausgehen. Ein Beispiel dafür ist die Decision Field Theory von Busemeyer & Townsend (Busemeyer & Townsend, 1993; Roe, Busemeyer, & Townsend, 1999), ein anderes Beispiel das Modell mit einem integrierten Schwellenwert, dem *desired level of confidence* (DLC) von Hausmann und Läge (2008). Hausmann und Läge gehen in ihrem Modell davon aus, dass Schritt für Schritt Informationen akkumuliert werden, bis ein individuelles Niveau an Urteilssicherheit erreicht wird.

Im Gegensatz zu den Modellen, welche von einer sequentiellen Verarbeitung von Informationen ausgehen, existieren, häufig in Konkurrenz, solche Modelle, welche eine parallele Verarbeitung von Informationen annehmen. Glöckner und Betsch postulieren zum Beispiel eine parallele Verarbeitung von Informationen und gehen von gespeicherten Assoziationen aus, wobei die mentale Repräsentation von Informationen als zeitlich aktives Netzwerk konzeptualisiert ist (Glöckner & Betsch, 2008). Ein weiteres Modell, das von einer parallelen Verarbeitung ausgeht, ist das exemplarbasierte Modell. Es verzichtet auf die Annahme, dass das Verhältnis zwischen einem Cue und einer gewissen Zielvariable im Gedächtnis gespeichert wird, sondern postuliert eine Speicherung von Merkmalen, welche in einer prototypischen Suche im Gedächtnis als „Exemplare“ wiedererkannt werden (z.B. Dougherty, Gettys, & Ogden, 1999; Juslin & Persson, 2002; Mitchell & Beach, 1990).

Ich konzentriere mich in dieser Arbeit auf die verschiedenen Gütemasse von probabilistischer Information bei der sequentiellen Informationssuche und teste ihre Möglichkeiten und Grenzen in verschiedenen Entscheidungssituationen. In einem ersten Teil stelle ich den aktuellen Forschungsstand zu den klassischen, probabilistischen Gütemassen im Rahmen des PMM dar und weise anschliessend auf die Lücken hin, die sich meines Erachtens daraus ergeben. Schlussfolgernd auf diesen Erkenntnissen werden die neuen, bzw. die Theorie erweiternden Gütemasse aufgezeigt, die wir entwickelt und experimentell überprüft haben, und die im Anschluss an diesen Überblick ausführlich dargestellt sind. Am Ende dieses Überblicks werden die Experimente diskutiert und mögliche zukünftige Forschungsideen aufgezeigt.

Forschungsstand zu der Cue-Güte von Informationen

Um ein Wissen über die Cue-Gütemasse und dessen Anwendung im prädeziationalen Suchprozess zu testen, wurden verschiedene aufwändige empirische Lernexperimente durchgeführt. Im Folgenden wird der aktuelle Forschungsstand zu den Lernexperimenten zu den Cue-Gütemasse Validität und Diskriminationsrate dargestellt.

Die Validität

Bröder (2000a) fokussierte auf die Entscheidungsregel der Heuristik, liess aber dennoch die Cue-Validitäten, welche für eine Kulturtechnik von extraterrestrischen Viltvodl'scher Völker sprachen, lernen. Die Versuchspersonen mussten jeweils auf eine von zwei Völkern tippen und das vermeintlich zivilisiertere auswählen, worauf sie eine Rückmeldung erhielten. Um das Validitätslernen zu optimieren, folgte auf jede Rückmeldung ein Feedback in der Form eines sich über die Lernphase verändernden Balkendiagramms, das die Trefferrate des jeweiligen Cues zusammenfasste. Bröder fand, dass die Versuchspersonen tendenziell hohe Validitäten unterschätzen und tiefe überschätzen. Numerisch waren die subjektiven und objektiven Validitäten nicht identisch, aber die Reihenfolge der subjektiven Validitäten entsprach der Reihenfolge der objektiven Validitäten. In einem dritten Experiment variierte Bröder die Lernmethode und unterschied zwischen „aktivem“ (wie oben beschrieben) und „passivem“ Lernen. Beim „passiven“ Lernen konnte nach 20 Runden „aktivem“ Lernen einfach die dynamische Entwicklung der Balkendiagramme beobachtet werden. Der Autor schlussfolgerte anhand seiner Ergebnisse, dass pragmatisch gesehen die vereinfachte „passive“ Lernmethode gerechtfertigt sei. Zudem folgerte Bröder, dass einzig eine lineare Gleichheit der subjektiven und objektiven Validitäten nötig wäre, um die sequentiellen Suchregeln gemäss PMM zu überprüfen und diese Linearität in diesen Experimenten gezeigt werden konnte. In dem dritten Experiment, dem Börsenspiel, ging Bröder deshalb dazu über, die Validitäten der Cues am Anfang des Experimentes numerisch vorzugeben und Informationskosten einzuführen. Da in diesen Studien die Cues nur in der Rangreihe der Cues gesucht werden konnte, machen die Befunde dieser Studie keine Aussage über die Anwendung der Suchregel, sondern bloss der Entscheidungsregel. Bei der Entscheidungsregel fand Bröder grosse individuelle Differenzen und schlägt deshalb einen idiographischen Ansatz für die Entscheidungsforschung vor.

Um aber das Lernen von Cue-Validitäten und das Etablieren von Cue-Hierarchien und deren Anwendung in der Suche zu prüfen, übernahmen Newell und Shanks (2003) das Börsenszenario von Bröder und adaptierten es soweit, dass die Cues nun sequentiell frei aufdeckbar waren und nicht wie bei Bröder nach vorgegebener Validitätsreihenfolge. In 60 Trainingsdurchgängen wurden den Versuchspersonen jeweils das ganze Cue-Muster (d.h. Cues mit binären Ausprägungen für zwei Aktien) vorgegeben. Auf Basis dieser Informationen mussten die Versuchspersonen danach auf eine Aktie tippen, worauf sie ein Feedback erhielten. Da die Personen im ersten Experiment die Validitäten sehr schlecht lernten, schalteten die Autoren in den weiteren Experimenten sogenannte hints (Tipps) ein. Das bedeutet, dass sie in der Mitte der Lernphase und direkt am Anfang der Testphase die korrekte Reihenfolge der Cue-Validitäten vorgaben. Genau gleich verfuhr Newell, Weston und Shanks (2003). Die „richtige“ Lösung der Cue-Rangreihe aber nach der Hälfte des Trainings und nach Abschluss des Trainings offen vorzugeben, macht das Cue-Lernen zur Alibi-Übung und auch die Befunde,

dass Personen die Informationen nach den vorgegebenen Rangreihen suchen, nicht weiter erstaunlich.

In den Experimenten von Bröder und Newell wurden die Validitäten immer aufgabenorientiert und in einem paarweisen Vergleich zum Lernen dargeboten. Das heisst, die Cue-Ausprägungen von den beiden zur Wahl stehenden Optionen wurden alle angezeigt und es musste anhand von diesen Informationen auf das zivilisiertere von zwei Völkern bzw. die bessere Aktie von zwei Aktien getippt werden. Dieses Vorgehen wurde gewählt, um die theoretischen Postulate des PMM zu überprüfen und sich dem klassischen Beispiel von Gigerenzer und Kollegen, dem Städteparadigma anzunähern. Wir stellen jedoch fest, dass in der natürlichen Umwelt Ausprägungshäufigkeiten nicht zwingend im Paarvergleich und aufgabenbezogen gelernt werden müssen. Aus diesem Grund liessen Hausmann, Christen & Läge (2005) Cue-Validitäten durch Beobachtung und nicht im Paarvergleich lernen. Das Lernen wurde in einem Millionärsquiz umgesetzt, in dem die Versuchspersonen beim Bearbeiten von 50 Wissensfragen gleichzeitig beobachten konnten, wie gut andere Personen (Telefonjoker) die richtige Lösung trafen. Den Versuchspersonen wurde nahegelegt, sich ein Bild von der Qualität der Antworten dieser Personen zu machen, da sie in der zweiten Phase des Experiments die Möglichkeit haben würde, gegen eine Gebühr eine oder mehrere dieser Personen um Rat zu fragen. Nach jeweils zehn Durchgängen wurden zudem die subjektiven Einschätzungen der Validitäten der vier personifizierten Cues (d.h. die Expertise der Joker) erhoben. In der anschliessenden Anwendungsphase waren die Personen beim Beantworten der schwierigen Quiz-Fragen auf sich alleine gestellt, konnten aber die Telefonjoker um ihren Tipp befragen. Es zeigte sich, dass die Validitäten der personifizierten Cues ordinal richtig eingeschätzt wurden. Gleichzeitig wurden jedoch die absoluten Validitäten systematisch und massiv unterschätzt. Methodologisch ist an der Arbeit von Hausmann et al. zu kritisieren, dass die Lernphase jeweils von den Einschätzungen der Validitäten unterbrochen wurde. Dieser Umstand lässt, wie die Autoren selbst bemerkten, offen ob die Einschätzung der Cue-Validität ein Fokus auf den momentanen Eindruck der Validität oder eher eine Beurteilung der Validität im gesamten Experiment ist.

Die berichteten Untersuchungen zeigen, dass die Lernbarkeit der Cue-Güte Validität nicht sehr einfach ist. Der in den meisten Experimenten verwendete Paarvergleich scheint nicht das ideale Setting zum Lernen von Cue-Gütemassen zu sein, sondern ein reines Häufigkeitslernen könnte dafür angebrachter sein. Wie die Validitäten in den verschiedenen Lernsettings jedoch wahrgenommen, gelernt und in der Informationssuche verwendet werden, ist eine bisher unbeantwortete Forschungsfrage.

Die Diskriminationsrate

Die bis hierhin dargestellten Experimente beschränkten sich auf das Lernen und Anwenden von Cue-Validitäten. Newell, Rakow, Weston und Shanks (2004) weiteten das Lernen auch auf die Diskriminationsrate aus. Sie liessen ihre Versuchspersonen die Validitäts- und Diskriminationsrate zusammen in 64 Durchgängen im bekannten Börsenszenario lernen. In der Lernphase waren demnach die Informationen von vier binären Cues vorhanden (teilweise diskriminierend und teilweise nicht) und es musste auf Aktie A oder B getippt werden, worauf die Personen dann ein Feedback über die Richtigkeit ihrer Antwort bekamen. Die Autoren

gingen dabei davon aus, dass in der Lernphase das kombinierte Mass von Validität und Diskriminationsrate, der Success, gelernt werden sollte. In der Testphase mussten dann die Informationen für gewisse Beträge gekauft werden, um eine Entscheidung zu treffen. Am Schluss des Experiments mussten die Versuchspersonen die Nützlichkeit der Cues¹, die Validitäts-, Diskriminations- und Successrate der Cues einschätzen. Dabei wurde die Validitätsrate unterschätzt und die Diskriminationsrate überschätzt. Das Fazit dieser Forschungsergebnisse war dann: Können die Personen die Reihenfolge der Cues frei wählen, dann setzt sich die Suchregel Success durch (siehe auch Lagnado, Newell, Kahan, & Shanks, 2006; Rakow, Newell, Fayers, & Hersby, 2005). In dem Experiment von Newell und Kollegen wurde Success aber als sequentielle Suchregel ohne Stoppen nach der ersten Information getestet, eher im Sinne von Hoffrage, Hertwig, & Czienskowski (2003), die die Usefulness als die Kombination von $V \cdot D$ definierten, ohne die Stoppregel zu spezifizieren. Die Anwendung von Success bzw. Usefulness wurde von Läge, Hausmann, Christen, & Daub (2005) in einem fiktiven chinesischen Städte-Experiment analog zum Städteparadigma überprüft. Die Autoren fanden, dass Success und Usefulness die „Sieger-Strategien“ gegenüber den Suchstrategien basierend auf der Validität bzw. der Diskriminationsrate sind, wobei Success eher bei beschränkter Informationssuche und Usefulness bei unbeschränkter Informationssuche angebracht ist.

In einem neueren Experiment verzichteten Christen, Hausmann und Läge (2006) wiederum auf das Lernen anhand eines paarweisen Vergleichs und liessen die Validität, die Diskriminationsrate und auch Success durch aktives Beobachten lernen. Dabei sollte den Versuchspersonen weder Hinweise über eine Cue-Rangreihe noch graphische Hilfen oder Erklärungen zur Cue-Güte gegeben werden. In dem Setting mussten die Personen anhand medizinischer Tests kranke Äffchen diagnostizieren und mit Medikamenten retten. Die Lernphase bestand aus 100-200 Durchgängen, wobei die Hinweise der fünf Tests gleichzeitig erschienen. Falls ein Test diskriminierte (wenn der Test nicht diskriminierte, zeigte er kein Resultat an), bestand das Testresultat in der binären Anzeige „+“ (Äffchen ist krank) oder „-“, (Äffchen ist gesund). Danach musste die Personen entscheiden, ob sie den Äffchen eine Medizin verabreichen wollten oder nicht. Eine Rückmeldung gab ihnen dann die Folge ihrer Entscheidung bekannt (Äffchen hätte überlebt oder wäre verstorben). In einer zweiten Phase mussten die Personen die Tests ganz spezifisch nach ihrer Validität, Diskriminationsrate oder Success-Rate einsetzen. Die Befunde zeigten, dass die Mehrzahl der Personen die Gütemasse in der Lernphase adäquat gelernt hatten und fähig waren, diese auch korrekt einzusetzen. Vor allem scheint die Diskriminationsrate lernbar zu sein und separiert von der Validität (und nicht nur in Kombination von Success oder Usefulness) auch korrekt angewendet zu werden.

Im Grossen und Ganzen weisen die bisherigen Befunde darauf hin, dass die experimentelle Lernbarkeit der Cue-Güte Validität und Diskriminationsrate nicht sehr einfach ist. Im speziellen zeigen die Befunde eine bessere Erlernbarkeit von Cue-Validität und Diskriminationsrate, wenn die Gütemasse statt im paarweisen Setting durch Beobachtung der Häufigkeiten gelernt werden. Diese Vermutung stützt sich jedoch auf ein paar wenige Experimente und weiterer Forschungsbedarf bleibt erhalten.

¹ Usefulness wurde bei den Autoren im Sinne von der allgemeinen Nützlichkeit eines Cues verwendet und mit Success gleichgesetzt.

Blinde Flecke

Bei der Darstellung des aktuellen Forschungsstandes zur Informationssuche von probabilistischen Informationen zeigen sich einige unerforschte, bildhaft gesprochen blinde Flecken im brunswikschen Linsenmodell. In Zusammenarbeit mit Damian Läge, Daniel Hausmann und Stephan Christen habe ich in den letzten Jahren im Rahmen des Nationalfonds-Projektes „Informationssuche beim Entscheiden“ (Beitrag Nr. 10-100282.01 und Nr. 10-107853.01) drei Schwerpunkte herausgearbeitet, welche unseres Erachtens bei den „klassischen“ Cue-Gütemassen Validität und Diskriminationsrate bisher vernachlässigt wurden.

Zufriedenheit mit der Informationslage

Es wird immer wieder diskutiert, ob Personen nun nach einer Information ihre Informationssuche beenden und eine nonkompensatorische Entscheidungsstrategie wählen oder ob sie nach mehreren Informationen suchen und diese kompensatorisch verarbeiten. Die Stoppregeln von Gigerenzer zu den Entscheidungsheuristiken *Take The Best*, *Take The Last* und *Minimalist* postulieren alle das Stoppen nach einer einzigen (diskriminierenden) Information. Sie sind folgendermassen definiert: „If one object has a positive cue value („1“) and the other does not (i.e. either „0“ or unknown value) then stop search and go on to [decision]“ (Gigerenzer et al., 1999, S.80). In unterschiedlichen experimentellen Studien war bisher jedoch zu beobachten, dass Personen, denen viel Information vorliegt, nur in seltenen Fällen auf den Versuch der vollständigen Aufdeckung aller Daten verzichten wollen und dann eher nach der mathematisch optimal aussehenden Wahl streben als nach einer Wahl aufgrund eingeschränkter Informationsverarbeitung (Bröder, 2000a, 2000b; Läge, Hausmann, & Christen, 2005). Dabei verletzen die Personen die Stoppregel der einfachen Entscheidungsheuristik und tendieren unter kognitiv anspruchsvollen Randbedingungen zum Scheitern. „[...] people are not always satisfied to make a decision on the basis of one piece of discriminating information, even when acquiring further information is potentially both financially and cognitively disadvantageous“ (Newell & Shanks, 2003, S.64). Das heisst, Personen stoppen ihre Informationssuche häufig zu spät. Dieses zu späte Stoppen kann unterschiedlich interpretiert werden. Aus einer ökonomischen Perspektive betrachtet, muss davon ausgegangen werden, dass die Leute gemäss einem Erwartung-mal-Wert Modell den Nutzen der Information falsch einschätzen, indem sie diese überschätzen. Von einem psychologischen Standpunkt aus gesehen wollen wir jedoch auf die Interpretationen des Fehlverhaltens verzichten und uns fragen, was der subjektive Zusatznutzen der gekauften Informationen ist, der zu einem verzögerten Stoppen führt. Verschiedene Schwellenwertmodelle erklären diesen Effekt damit, dass Menschen ein gewisses Anspruchsniveau an Urteilssicherheit bei der Informationssuche erreichen wollen (Hausmann-Thürig, 2004; Hausmann & Läge, 2005a) oder eine Konfidenzschwelle (Schoppek & Putz-Osterloh, 2004) haben. Im Modell von Hausmann und Läge wird davon ausgegangen, dass Personen solange Informationen suchen, bis die Informationen ihr *Desired Level of Confidence* (DLC) decken. Die experimentellen Resultate zeigen, dass die Annahme eines personspezifischen DLCs die Beendigung von Informationssuche gut vorhersagen kann. Wenn nun aber nach einer Information die Suche nicht gestoppt wird, was in ca. der Hälfte aller Fälle vorkommt, stellt sich die Frage, anhand welcher Kriterien die Personen dann weitersuchen. Dabei interessiert uns vorallem die Abhängigkeit des Stoppens von der Informati-

onskonsistenz, d.h. von zwei sich gegenseitig bestätigenden bzw. widersprechenden Hinweisen. Wir untersuchen daher die Möglichkeit, dass sich der kritische Punkt des Stoppens als ein erreichter Grad an Zufriedenheit mit der Informationslage beschreiben lässt. Diese Urteilszufriedenheit kann sich aus mehreren Komponenten zusammensetzen, von denen wir annehmen, dass die Urteilssicherheit und die Informationskonsistenz zwei der zentralen Faktoren sind.

Lernumwelten, in denen „keine“ Information zu einer Information wird

Im theoretischen Rahmen der Urteilsheuristiken wird davon ausgegangen, dass die Cue-Informationen gemäss der gelernten Validität lexikografisch gesucht werden. Um die ökologische Validität eines Cues aber bestimmen zu können, braucht es eine Referenzgruppe von Objekten, bei welchen dieser Cue diskriminiert. In Umwelten, in welchen die Cues verschiedene Diskriminationsraten haben, muss theoretisch für jeden Cue eine eigene Referenzmenge mit unterschiedlicher Grösse konstruiert und mit den anderen verglichen werden, um eine Hierarchie der Cue-Validitäten bilden zu können. Newell und Kollegen kritisieren deshalb, dass diese Aufgabe, eine Hierarchie basierend auf verschiedenen Häufigkeiten von Inferenzen zu etablieren, sehr schwierig zu sein scheint (Newell et al., 2004). Basierend auf den Erkenntnissen aus den Experimenten zur Quiz-Show und dem Äffchentest nehmen wir an, dass für die Konstruktion der Rangreihen (sowohl der Validitäten, wie auch der Diskriminationsrate) diese Paarvergleiche der Referenzmengen nicht benötigt werden, sondern ein einfacher „innerer“ Zähler von Häufigkeiten genügt. Denn im Äffchen-Experiment wurde keine Wahl zwischen zwei Alternativen erzwungen, sondern es musste jeweils entschieden werden, ob einem Äffchen ein Medikament verabreicht werden soll oder nicht. Dementsprechend ist die Diskriminationsrate dadurch definiert, ob ein Test einen Hinweis liefert oder nicht und nicht wie bis anhin, ob ein Cue zwei gleiche Informationen zu den zwei zur Verfügung stehenden Alternativen angibt oder zwei verschiedene. Da im Äffchen-Setting die Cue-Gütemasse gut gelernt werden, nehmen wir an, dass Personen nicht die Diskriminationsrate lernen, sondern die sogenannte Ausprägungsrate. Diese Ausprägungsrate definieren wir durch die Häufigkeit von Merkmalen, d.h. durch die Häufigkeit von positiven Ausprägungen eines Merkmales innerhalb einer Referenzgruppe. Wir postulieren, dass die Ausprägungsrate mit einem „inneren Zähler“ einfacher gelernt werden kann als die Diskriminationsrate. Cue-Validitäten basieren theoretisch auf der Idee eines inneren Häufigkeiten-Zählers, wie dieser von Hasher und Zacks postuliert wurde (1979). Dougherty, Franco-Watkins, & Thomas (2008) kritisieren zwar, dass die Annahme eines automatischen Häufigkeiten-Zählers einerseits durch aktuelle Literatur nicht gestützt wird und andererseits das Problem besteht, dass ein solcher Zähler, um die ökologische Validität berechnen zu können, neben dem Wissen übers Vorkommen von Merkmalen auch Wissen übers Nicht-Vorkommen dieser Merkmale benötigt. Dieser Einwand mag zwar gerechtfertigt sein, dennoch gehen wir davon aus, dass ein „innerer“ Zähler ein einfacheres Erklärungsmodell ist als ein ständiger Paarvergleich von vorhandenen bzw. nicht vorhandenen Informationen, und untersuchen diese beiden Konzepte in verschiedenen Lernexperimenten deshalb genauer.

Urteilsheuristiken legen zudem fixe Such- und Stoppregeln nahe, anhand derer Informationsquellen sequentiell befragt werden sollten. Über diese festgelegten Suchreihenfolgen (z.B. in absteigender Validitätsfolge wie bei *Take The Best*) hinausgehend postulieren wir nun eine adaptive sequentielle Informationssuche in Abhängigkeit der Qualität von nicht diskriminie-

render Information (positiv / positiv versus negativ / negativ). Wenn beide zur Verfügung stehenden Optionen über dieselbe Cue-Ausprägungen verfügen, ist dies als Zusatzinformation zu werten und legt die weitere Suchrichtung fest. Nach zwei positiven Cue-Ausprägungen (+ +) wird als nächstes ein Cue mit niedriger Ausprägungsrate aufgesucht, bei zwei negativen Cue-Ausprägungen (- -) wird ein Cue mit hoher Ausprägungsrate gewählt. Somit postulieren wir eine adaptive Informationssuche in Abhängigkeit der nicht-diskriminierenden Information und in Richtung der entsprechenden Ausprägungsrate.

Adaptives Verhalten unter Einbezug einer Zeitdimension

Im Bereich des menschlichen Handelns und Entscheidens wird unter Information ein Wissen verstanden, dem in der jeweiligen aktuellen Situation eine gewisse Bedeutung beigemessen wird. In diesem Zusammenhang dient die Information zur Erhöhung der Sicherheit, die richtige Entscheidung zu einem gewissen Zeitpunkt zu treffen. Nicht nur, dass es gute und weniger gute Zeitpunkte für eine Entscheidung gibt (vgl. Kerstholt & Raaijmakers, 1997), sondern auch die Informationen, welche einer Entscheidung zugrunde liegen, ändern sich über die Zeit hinweg. Deswegen sollte die Suche nach den probabilistischen Hinweisen auf die beste wählbare Alternative auch möglichst eng an den Zeitpunkt der Entscheidung selbst gekoppelt sein, wenn man eine maximale Validität von Prädiktorvariablen (Cues) für eine Zielvariable (Entscheidungsalternative) erreichen will. Die bisherige experimentelle Entscheidungsforschung hat, soweit sie sich überhaupt mit Informationssuche als einem aktiven Prozess zur Vorbereitung einer Entscheidung beschäftigt hat, die Zeitdimension dadurch ausgeschlossen, indem die Informationssuche und Entscheidung auf denselben Zeitpunkt gelegt wurden (Bröder, 2000a, 2000b, 2002, 2003, 2005; Bröder & Schiffer, 2003, 2006; Gigerenzer, Todd, & ABC Research Group, 1999; Hausmann & Läge, 2005b; Läge, Hausmann, & Christen, 2005; Läge, Hausmann, Christen et al., 2005; Newell et al., 2004; Newell & Shanks, 2003; Newell et al., 2003; Payne, 1976; Payne & Bettman, 2000; Payne, Bettman, & Johnson, 1988; Payne, Bettman, & Johnson, 1993; Rieskamp & Hoffrage, 1999). Die dabei eingesetzten Information Boards waren dann entsprechend so strukturiert, dass es sich wiederholende Durchgänge gab, in denen jeweils Informationen gesucht und auf der Basis dieser Informationen entschieden werden konnte. Der nächste Durchgang hatte dann (von der Information her gesehen) nichts mehr mit dem vorausgehenden Durchgang zu tun. Dabei wurden diese Durchgänge als unabhängige Einzelereignisse betrachtet, d.h. die Versuchspersonen konnten keine Informationen aus vorhergehenden Durchgängen für die aktuelle Entscheidung nutzen, noch konnten sie zum aktuellen Zeitpunkt bereits Informationen für zukünftige Entscheidungen suchen. Entscheiden spielt sich aber im Alltag in der Regel im Rahmen von Zeitverläufen ab. Möglicherweise relevante Information stammt dabei oft aus der Vergangenheit, und es muss eingeschätzt werden, ob sich eine Aktualisierung lohnt. In vielen Fällen (z.B. bei Alltagsentscheidungen wie etwa dem Kauf eines Neuwagens oder bei der Einstellung eines neuen Mitarbeiters, etc.) scheint das statische Modell in seinen Grundzügen anwendbar zu sein, in anderen hingegen, und dazu zählt beispielsweise der Aktienmarkt, führt die Annahme stabiler Umwelten zu völlig unbrauchbaren Verhaltensvorhersagen. Hier ist vielmehr eine Erweiterung auf sich verändernde Umwelten einerseits und stabile/instabile Informationen andererseits zu leisten. Um der Veränderbarkeit von Information Rechnung zu tragen, führen wir

somit eine Stabilitätsrate in Bezug auf Prädiktorvariablen (Cues) und Zielvariable (Entscheidungsalternative) ein.

Ein Grundsatz der „Adaptiven Toolbox“ ist die Adaptivität an die verschiedenen statistischen Umwelten. Inwieweit ein solches adaptives Verhalten bei einem Wechsel von einer statischen zu einer zeitlich dynamischen Umwelt beobachtbar ist, soll des Weiteren untersucht werden.

Um die dargestellten und bisher vernachlässigten Lücken in der aktuellen Forschung bzw. Theorie etwas zu verkleinern, schlagen wir deshalb verschiedene konzepterweiternde Gütemasse vor: die Informationskonsistenz, die Cue-Ausprägungsrate und die Erweiterung der Entscheidungssituation um eine Zeitdimension.

Beiträge zum aktuellen Forschungsstand

Die Informationskonsistenz

Wir postulieren, dass die Informationskonsistenz, d.h. ob zwei Informationen sich widersprechen oder bestätigen, ein Bestandteil der Urteilszufriedenheit ist und damit einen Einfluss auf die Informationssuche und das Stoppen hat. Unter dieser Annahme, dass unterschiedliche Informationskonsistenz bei konstanter Urteilssicherheit die Urteilszufriedenheit moderiert und damit den Zeitpunkt des Stoppens der Informationssuche beeinflusst, führten wir in einem Pferderenn-Szenario eine erste Phase durch, in welcher Cues aktiv gesucht werden mussten. Dabei liessen wir diese Cues hinsichtlich der subjektiven Urteilssicherheit und Urteilszufriedenheit von den Versuchspersonen bewerten und erhielten so gleich eine Schätzung des individuellen DLC von jeder Versuchsperson. In einer zweiten Phase dieses Experiments wurden die Probanden mit denselben Cue-Validitäten konfrontiert, jedoch diesmal nicht als einzelne Cues, sondern als Kombination von zwei (mal bestätigenden, mal widersprüchlichen) Cues. In **Kapitel 1 (Wie zufrieden machen Informationen? Die Bewertung von Informationen in Bezug auf Urteilssicherheit und Urteilszufriedenheit)** wird das Konzept der Urteilszufriedenheit mit jener der Urteilssicherheit verglichen und der Einfluss der Informationskonsistenz explorativ erforscht. Wir finden einen Unterschied in der A-priori-Validität von zwei Informationen und der A-posteriori-Validität. Die A-priori-Validität ist dann heranzuziehen, wenn es um die Informationssuche geht und die Frage, welche Informationsquelle die Personen bevorzugen, wenn sie noch nicht wissen, ob die beiden Informationen bestätigend oder widersprüchlich sein werden. Die A-posteriori-Validität ist wichtig, wenn es um den Vergleich von widersprüchlicher und bestätigender Information und deren Zusammenhang zur subjektiven Urteilssicherheit und Urteilszufriedenheit geht. Die gefundenen Resultate erhärten die Idee, dass nicht nur die Validität einer Information die Informationssuche beeinflusst, sondern auch andere Faktoren, wie die Konfliktfreiheit von zwei Informationen oder aber wie viele Informationen vorliegen. Wir können daher von einer adaptiven Informationssuche und adaptivem Stoppen ausgehen. Anfangs spielt die Validität eine Rolle. Wird dann das Anspruchsniveau an Urteilssicherheit nicht erreicht und deshalb weitergesucht, kommt die Urteilszufriedenheit ins Spiel. Obwohl die Urteilszufriedenheit eng mit der Urteilssicherheit verbunden ist, zeigen die Versuchspersonen vor allem bei der Integration von zwei Informationen spezifische Unterschiede in der Bewertung der beiden Konzepte. Die Informationskonsistenz, also ob zwei Informationen widersprüchlich sind oder nicht, und die Anzahl der

Informationen können als Variablen des Konzeptes der Urteilszufriedenheit verstanden werden. Unsere Befunde zeigen, dass das Konzept der Urteilszufriedenheit einen Aspekt in der prädeziSIONalen Phase der Entscheidung abdeckt, welcher von der Validität der Information und der subjektiven Urteilssicherheit nicht erklärt werden kann.

Lernumwelten und die Cue-Ausprägungsrate

Wir gehen einerseits davon aus, dass die Cue-Gütemasse nicht gemäss einem Paarvergleich gelernt werden, sondern eher durch Beobachtung der Häufigkeiten und andererseits, dass die Informationssuche nicht streng sequentiell, sondern adaptiv an die vorhandene Informationslage (je nach Ausprägungsrate der Cues) angepasst wird.

In **Kapitel 2 (Experiment 1 und 2)** untersuchen wir anhand des fiktiven chinesischen Städteparadigmas (analog zu Läge, Hausmann, Christen et al., 2005) die sequentielle Informationssuche in Abhängigkeit dieser Ausprägungsrate. Die Resultate zeigen, dass Personen die Informationssuche adaptiv an die vorhandene Informationslage anpassen: Nicht diskriminierende Information (z.B. beide Städte haben keine Tageszeitung) scheint den Personen einen Anhaltspunkt über die Grössenordnung der Städte zu geben, so dass sie in der Folge bevorzugt Cues auswählen, die innerhalb der vermuteten Referenzklasse eine hohe Diskriminationsleistung besitzen. In **Kapitel 2 (Experiment 2 und 3)** untersuchen wir, auf welche Art die Validität und Diskriminationsrate gelernt und dann in der Informationssuche angewendet werden. Wir finden, dass die Ausprägungsrate einfach zu erlernen ist und dass anhand der Ausprägungsrate auf die Diskriminationsrate geschlossen werden kann. Diese Befunde unterstreichen somit experimentell den Nutzen der theoretischen Annahme eines „inneren“ Zählers. Um die Validität zu erlernen, scheint es jedoch den Paarvergleich oder zumindest eine Aufgabe mit direktem Feedback zu benötigen.

Zeitlich dynamische Umwelten und Adaptivität

Dem Aspekt von einer sich ständig im Fluss befindliche Informationslandschaft wird in den Kapiteln 3 bis 6 vertieft anhand von verschiedenen Schwerpunkten Beachtung geschenkt. In **Kapitel 3 (Entscheiden in zeitlich dynamischen Umwelten: Suchen und Stoppen mit unterschiedlich stabilen Cues)** werden im Rahmen des bekannten Börsenszenarios neben der Einführung der Zeitachse eine Stabilitätsrate a) der Zielvariablen und b) der Prädiktorvariablen definiert. Mit einem Fokus auf die Güte von aktuellen und vergangenen Informationen werden die Such- und Stopppregel in einerseits einer statischen Umwelt (mit statistisch unabhängigen Ereignissen) und andererseits einer dynamischen Umwelt (mit zeitlich verbundenen Ereignissen) analysiert. Es zeigt sich, dass Versuchspersonen in einer zeitlich dynamischen Umwelt mit weniger Information auszukommen glauben, dass sie teilweise ganz ohne neue Informationssuche entscheiden und dass sie Cues mit hoher Stabilität (die also für zukünftige Entscheidungen noch verwendbar sind) in stärkerem Maße berücksichtigen als in einer Vergleichsbedingung ohne Zeitachse. In einer bedeutenden Zahl von Fällen ziehen die Versuchspersonen die längerfristig stabile Information bei der Informationssuche sogar der momentan valideren Information vor. Interessant ist, dass die beschriebenen Verhaltensänderungen sich vornehmlich auf ein Drittel der Personen konzentrieren.

Wenn Personen den vergangenen Informationen aber nicht vertrauen und ihre Urteilssicherheit erhöhen möchten, dann kann eine erneute Informationssuche dabei im Konflikt mit anderen Aktivitäten stehen, so dass die Ressource Zeit eine Rolle für die Frage der Informationssuche beim Entscheiden spielt. Um diese alltägliche Situation für die Entscheidungspsychologie erforschbar zu machen, entwickelten wir ein Trinkwasser-Szenario, welches eine entsprechende zeitlich dynamische Umwelt darstellt. In **Kapitel 4 (Die zeitliche Dynamik von Umwelten: Heuristisches Wahlverhalten am Beispiel eines Trinkwasser-Szenarios mit und ohne Zeitachse)** spielt neben der Ressource Zeit vor allem auch das Wissen um die zeitliche Stabilität der wählbaren Alternativen eine zentrale Rolle. In einem ersten Experiment mit diesem Szenario wurde überprüft, wie weit die Versuchspersonen auf eine solche zeitlich dynamische Umwelt reagieren (im direkten Vergleich zu einer statischen Umwelt mit Einzelentscheidungen). Die Befunde bestätigen im Wesentlichen die aufgestellten Hypothesen hinsichtlich Suchregel, Stoppregel und Entscheidungsregel der verwendbaren sequentiellen Entscheidungsheuristiken. Sie zeigen aber auch, dass zumindest ein Teil der Probanden nicht oder nur sehr wenig auf den Wechsel der statistischen zur zeitlich dynamischen Umwelt reagiert und daher wenig adaptiv sind.

Weil die für die Informationssuche aufzubringende Zeit ein wichtiger Kostenfaktor ist, der Umfang und Qualität derjenigen Information mitdefiniert, die zum Zeitpunkt der Entscheidung zur Verfügung steht, wird diese wichtige Alltagsressource in dem Trinkwasser-Szenario der entscheidungspsychologischen Forschung zugänglich gemacht. Die strukturelle Parallelität des Trinkwasser-Szenarios mit dem Börsenszenario erlaubt es, das Verhalten bei der Informationssuche mit Zeit (Suchregel, Stoppregel und Risikoappetenz in der Entscheidung) direkt mit der traditionell beforschten Ressource Geld in **Kapitel 5 (Ressource Zeit statt Geld: Ein Vergleich zweier Szenarien zur Entscheidungsforschung)** zu vergleichen. Es zeigt sich, dass die Versuchspersonen im Wesentlichen bei der Ressource Zeit ein ähnliches Verhaltensmuster im Umgang mit Information zeigen wie bei der Ressource Geld. Jedoch sind deutliche Tendenzen festzustellen, bei der Ressource Zeit eher mehr Informationen zu suchen und risikoreicher zu entscheiden. Offenbar wird mit der Ressource Zeit grosszügiger umgegangen als mit der Ressource Geld.

Das adaptive Verhalten ist einer der Grundsätze des theoretischen Rahmens der Adaptiven Toolbox. Einerseits stellt sich die Frage nach dem adaptiven Verhalten im Bezug auf diese unterschiedlichen statistischen Umwelten und andererseits nach der Adaptivität beim Wechsel von einer Umwelt in eine andere statistische Umwelt. In **Kapitel 6 (Adaptivität der Informationssuche in statischen und in dynamischen Umwelten)** wird diese Adaptivität anhand des Trinkwasser-Szenarios in einem 2x2 Design untersucht, wobei ein Faktor die statische Umwelt ist, also einerseits die Umwelt von Einzelereignissen und andererseits die dynamische Umwelt mit Zeitdimension und der zweite Faktor die Position ist, in welcher diese statistische Umwelt jeweils an erster bzw. zweiter Stelle im Experiment bearbeitet wird. In Bezug auf die abhängige Variable des Stoppverhaltens zeigen die Personen individuelle Präferenzen, welche nur in geringem Ausmass an die statistische Umwelt angepasst werden. Das Entscheiden hingegen ist sehr adaptiv und wir finden einen Einfluss der in der Vergangenheit getroffenen Wahl. Die Personen verhalten sich mehr oder weniger adaptiv, wenn sie mit einer

statistischen Umwelt konfrontiert werden, bleiben jedoch in dem dort angewendeten Verhalten nach einem Wechsel in eine andere statistische Umwelt verhaftet.

Kritische Bemerkungen

Da verschiedene Lernexperimente gezeigt haben, dass das Erlernen der Reihenfolge der verschiedenen Cue-Gütemasse schwierig und sehr aufwändig ist, verwendet unsere und andere Forschungsgruppen (z.B. Rieskamp & Hoffrage, 1999; Rieskamp & Hoffrage, 2008; Rieskamp & Otto, 2006) seit geraumer Zeit Experimentalsettings, in denen die Gütemasse der Cue-Informationen nicht mehr erlernt werden, sondern den Versuchspersonen direkt als Zahlen auf dem Information Board vorgegeben werden. Sowohl beim Börsenszenario wie auch beim Trinkwasser-Szenario und dem Pferderennen gaben wir die Cue-Gütemasse ebenfalls jeweils numerisch vor. Gigerenzer et al. (1991) kritisierte dieses Vorgehen mit dem Hinweis, dass die Informationssuche in dem PMM Modell als ein Informationsabruf im Gedächtnis gemäss den gelernten Cue-Informationen konzipiert sei. Inwieweit dieses Vorgehen der Vorgabe der Cue-Gütemasse sinnvoll ist, wird anhand der Anwendung und deren Gültigkeit für die Gütemasse Validität, Diskriminationsrate und Stabilitätsrate in **Kapitel 7 (Präsentation von Cue-Gütemassen im Information Board Setting: Validität, Diskriminationsrate und Stabilität)** dargestellt.

Um die nun folgende Diskussion der verschiedenen Beiträge nachzuvollziehen, müsste die interessierte Leserin bzw. der interessierte Leser an dieser Stelle die zuvor dargestellten sieben Kapitel auf den Seiten 22 bis 153 lesen.

Diskussion

Die Informationskonsistenz

Mit dem Experiment zur Informationskonsistenz postulierten wir als Punkt des Stoppens eine Zufriedenheit mit der Informationslage. Die Resultate zeigen, dass anfangs die Validität und damit einhergehend das Anspruchsniveau an Urteilssicherheit eine Rolle spielt und das Stoppen der Personen bestimmt. Betreiben die Personen aber kein *One Reason Decision Making*, sondern suchen weiter, dann kommt die Urteilszufriedenheit ins Spiel. Obwohl die Urteilszufriedenheit eng mit der Urteilssicherheit verbunden ist, zeigen die Versuchspersonen vor allem bei der Integration von zwei Informationen spezifische Unterschiede zur Urteilssicherheit auf. Die Informationskonsistenz, also ob zwei Informationen widersprüchlich sind oder nicht, und die Anzahl der Informationen können als Variablen des Konzeptes der Urteilszufriedenheit verstanden werden. Das Konzept der Urteilszufriedenheit scheint somit einen Aspekt in der prädezisionalen Phase der Entscheidung abzudecken, welcher von der Validität der Information und der subjektiven Urteilssicherheit nicht erklärt werden kann. Genauer zu untersuchen bleibt unter anderem die Interaktion der Validitäten von mehreren Cues mit der Urteilszufriedenheit.

Es kann anhand dieses Experimentes nicht eindeutig darauf geschlossen werden, ob die Versuchspersonen die Suche des Expertenpaares bloss auf die A-priori-Validität stützen und somit eindeutig sequentiell die Informationen suchen oder die beiden Informationen integrieren und eine A-posteriori-Validität einbeziehen. Es könnte durchwegs sein, dass auch die A-

posteriori-Validität unter der Annahme von widerspruchsfreien Informationen eine Rolle spielt, also die zwei Validitäten so zu einer Gesamtvalidität verrechnet werden. Für ein weiteres Experiment müssten nun die Informationspaare so geschaffen sein, dass beide Informationsquellen, wenn sie widerspruchsfreie Informationen abgäben, dieselbe A-posteriori-Validität hätten, aber das eine Paar klar eine sequentielle und streng lexikografische Suche, das andere Paar hingegen eine parallele Integration der beiden Informationen vorhersagen würde. Mit einem solchen Setting könnte man die methodische Kritik, dass die Erhebung anhand des Information Boards eine sequentielle Verarbeitung begünstigt, zu einem gewissen Teil abschwächen. Die Kritik, dass die verdeckten Boxen, welche mit einem gewissen Geld- oder Zeitverlust aufgedeckt werden müssen, eine sequentielle Verarbeitung begünstigen könnten und die Resultate bloss zu Artefakten des Settings machten (z.B. Glöckner, 2006), hat seine Berechtigung, wenn man von bewussten auf unbewusste Prozesse schliessen will. Meines Erachtens können aber jegliche Resultate zu den bewussten Prozessen nicht als Manifest der sequentiellen Verarbeitung von unbewussten Prozessen interpretiert werden, sondern müssen in dem methodischen Kontext differenziert und mit Verzicht der Generalisierung auf allgemeine Wirkmechanismen betrachtet werden. Nichtsdestotrotz fanden Bröder und Gaissmaier (2007) in einer Analyse von Reaktionszeiten bei sechs Gedächtnisexperimenten, dass die Annahme einer sequentiellen Informationssuche bestätigt werden kann. Weitere Forschung sollte daher differenzierter zeigen, in welchen natürlichen und experimentellen Situationen parallele bzw. sequentielle Verarbeitungsschritte auf welche Weise bewusst und unbewusst interagieren oder eben nicht.

Lernumwelten und Cue-Ausprägungsrate

Mit dem Konzept der Ausprägungsrate kann die Theoriebildung zur Validität und Diskriminationsrate genauer spezifiziert werden: Die Ausprägungsrate ist eine intuitiv leicht erlernbare Güte eines Hinweis-Cues. Doch bevor eine theoretische Redefinition der Validität vorzunehmen ist, sind weitere Aspekte derer Erlernbarkeit abzuklären. Die Personen, die die Gütemasse im paarweisen Setting erlernen, lernen diese Validität nämlich trotz allem noch besser als die Personen, welche die einfachen Häufigkeiten lernen. Fraglich ist dabei, ob es tatsächlich der Paarvergleich ist, welcher notwendig ist für diese Lernleistung, oder ob es nicht vielmehr das aufgabenorientierte Lernen anhand eines Feedbacks ist. Dazu müsste in einem weiteren Experiment das Ausprägungsrate-Lernen ebenfalls aktiver (z.B. durch Tippen auf die Grösse einer Stadt) und mit einem Feedback über die Richtigkeit dieser Einschätzung gestaltet werden. Enkvist, Newell, Juslin, & Olsson (2006) fanden nämlich, dass ein aktives Lernen (d.h. sie gaben das Kriterium vor und die Merkmale der Cues mussten aktiv ausgewählt werden) je nach Situation zu besserer Leistung führte und Pachur und Olsson (2008) fanden unterschiedliche Strategiewahl, je nach Art des Feedbacks (d.h. ob es absolut oder komparativ war). Solche Unterschiede in der Aufgabenstellung und im Feedback könnten auch beim Lernen und Anwenden der Validität einen Einfluss haben. Obwohl Bröder beim Erlernen der Validitäten keinen Unterschied zwischen aktivem und passivem Lernen fand (Bröder, 2000a), sollte zum Zwecke der Vergleichbarkeit die Aufgabenstellung der beiden Lernformate ähnlicher gestaltet werden.

Bisher wurde angenommen, dass, um die ökologische Validität eines Cues zu bestimmen, es eine Referenzgruppe von Objekten benötigt, bei welchen dieser Cue diskriminiert. Verschie-

denen Autoren haben bisher die ökologische Validität genommen, um die Suchregel von z.B. *Take The Best* zu untersuchen (z.B. Todd & Dieckmann, 2005). Dieckmann und Todd stellen fest (2004), dass die ökologische Validität der Cues durch gespeicherte oder laufende Erfahrungen berechnet werden muss. Subjektive Cue-Hierarchien korrelieren zwar mit der ökologischen Cue-Validität, aber es sei nicht dasselbe, weil die Absenz von Information genau gleich bewertet werden müsste wie das Vorhandensein, kritisieren Dougherty, Franco-Watkins, & Thomas (2008). Gigerenzer und Kollegen präzisieren heute, dass die Cue-Validitäten einer subjektiven Rangreihe entsprächen und nicht einer ökologischen Validität (Gigerenzer, Hoffrage, & Goldstein, 2008). Newell und Kollegen kritisieren, dass diese Aufgabe, eine Hierarchie basierend auf verschiedenen Häufigkeiten von Inferenzen zu etablieren, sehr schwierig zu sein scheint (Newell et al., 2004). Wir stellen fest, dass man für die Konstruktion der Rangreihen, sowohl der Validitäten wie auch der Diskriminationsrate, diese Paarvergleiche der Referenzmengen nicht benötigt, sondern ein einfacher „innerer“ Zähler genügen könnte. Weitere Forschung sollte aufzeigen, wie diese Ausprägungsrate mit der Validität beim Lernen interagiert und ob sie z.B. auch bei der Rekognitionsheuristik (Goldstein & Gigerenzer, 2002), im Sinne von einer Abhängigkeit der weiteren Informationssuche (je nach dem, ob zwei Städte wiedererkannt werden oder beide Städte nicht erkannt werden) eine Rolle spielt.

Unsere Befunde zu den gelernten Rangreihen der Cue-Gütemasse zeigen grosse individuelle Differenzen. Deshalb ist es notwendig neben den Auswertungen auf Mittelwertsebene auch die individuellen Daten auszuwerten. Da aber nicht nur zwischen den Personen grosse Unterschiede im Lernen wie auch der Anwendung von Cue-Gütemassen bestehen, sondern auch innerhalb einer Person (eine Person mag sich konsistenter in der Informationssuche verhalten als eine andere), genügt meines Erachtens der von Bröder vorgeschlagene idiographischer Ansatz zur Klassifikation der Entscheidung (eine Klassifikation der Entscheidungsstrategie anhand einer Regressionsanalyse) nicht. Um dieser Tatsache der intraindividuellen Differenzen bei der Informationssuche ebenfalls Rechnung zu tragen, könnte z.B. die Strategieklassifikation der Informationssuche anhand eines Quotienten der Strategiekonstanz differenzierter beschrieben werden.

Wie subjektive Hierarchien von Cue-Validitäten erstellt werden und welche Erfahrungen zur Repräsentation von welcher Validität und Diskriminationsrate führen, konnte in dieser Studie nicht abschliessend beantwortet werden. Gezeigt haben sollte die vorgelegte Studie aber eine Erweiterung in der Theoriebildung zu den Arten des Lernens von Cue-Gütemassen und zu der Adaptivität in Bezug auf den Prozess der Informationssuche. Weiterführende Forschung könnte wichtige Erkenntnisse über die Spezifität des theoretischen PMM in Bezug auf aufgabenorientiertes bzw. beobachtungsorientiertes Lernen von Cue-Gütemassen erbringen.

Zeitlich dynamische Umwelten und Adaptivität

Mit den Erkenntnissen aus dem um die Zeitdimension erweiterten Börsenszenario ist ein neues Feld definiert, welches bei der Berücksichtigung zeitlich dynamischer Umwelten von der Entscheidungsforschung noch weiter zu bearbeiten ist. Die Menschen reagieren sensitiv auf solche zeitlich dynamischen Umwelten: Zumindest einige suchen nicht nur gegenwartsorientiert Informationen, sondern berücksichtigen auch die Vergangenheit und teilweise sogar die Zukunft. Auch im Trinkwasser-Szenario beziehen die Personen vergangene Informationen

mit in die Planung der anstehenden Informationssuche ein und optimieren so ihr Suchverhalten in Bezug auf das Kosten-Nutzen-Verhältnis. Die beiden Szenarien haben somit zu einer Reihe von ersten Erkenntnissen zur Implementierung einer Zeitachse, und das Trinkwasser-Szenario ins besonders zur Implementierung der Ressource Zeit im Informationssuch- und Entscheidungsverhalten geführt. Die mit der Zeitdimension einher gehende Stabilität wird zum Zeitsparen genutzt und erbringt damit einen prospektiven Gewinn. Nicht oder praktisch nicht genutzt wird die Stabilitätsinformation hingegen im Hinblick auf die Nutzung der Information für das Abschätzen von zukünftigen Ereignissen. Es wird deswegen für zukünftige Forschung spannend sein, eine grundsätzliche Asymmetrie der Informationsnutzung im Hinblick auf Vergangenheitsnutzung und Zukunftsplanung zu überprüfen und, sofern sie sich erhärten sollte, ihren psychologischen Gründen nachzuforschen.

Der unterschiedlichen Charakteristik von Zeit und Geld im Hinblick auf Kumulierbarkeit und Einsatz dieser Ressourcen ist schwierig Rechnung zu tragen: Zeit ist eine Ressource, die kontinuierlich vergeht und selten als unmittelbarer Nutzen im Fokus der Entscheidung steht (jeweils im Unterschied zu Geld). Das gilt es auch im Experiment zu berücksichtigen, und uns scheint dies mit dem Trinkwasser-Szenario im Grundzug gelungen zu sein. Allerdings ergeben sich dadurch – betrachtet aus dem rein numerischen Blickwinkel – etwas unterschiedliche Kosten-Nutzen-Erwartungen bei den Entscheidungen. Sie auszuräumen, würde ggf. eine leichte Anpassung des Börsenszenarios bedeuten, z.B. indem man einen konstanten Betrag von 10'000 Franken pro Durchgang in einen Optionsschein investiert und die Informationskosten erst im Nachhinein abgezogen werden. Oder beim Trinkwasser-Szenario die asymmetrisch aufgebaute Kosten-Nutzen-Struktur (das Wegschütten von Wasser benötigte keinerlei Ressourcen) auszulösen, indem für zukünftige Experimente mit diesem Trinkwasser-Design auch für das Wegschütten einer Wasserladung jeweils einen Aufwand (z.B. 1 Stunde pro Lieferung) angesetzt wird. Es gibt jedoch auch subtilere Unterschiede in den beiden Ressourcen, die fundamentalerer Natur sind und sich durch Anpassungen in den Szenarien nicht ausräumen lassen: Beispielsweise erscheint es uns plausibel anzunehmen, dass das Erleben von Misserfolg bei beiden Ressourcen unterschiedlich sein könnte: Beide Szenarien-Typen charakterisieren sich durch unterschiedliche Kostenstrukturen bezüglich der "False Alarms" und der "Missed Hits". Diese strukturellen Unterschiede zwischen der Ressource Zeit und der Ressource Geld scheinen uns in der Sache selbst zu liegen und sind durch Parallelisierung im Experiment nicht oder nur unzureichend aufzuheben. Deshalb mag es nicht nur in diesem Experiment, sondern generell schwierig sein zu beurteilen, ob die beobachteten Verhaltensunterschiede allein auf die unterschiedlichen Ressourcen in den beiden Szenarien oder auch auf Unterschiede zwischen den Kosten-Nutzen-Strukturen zurückzuführen sind. Die Beantwortung dieser Fragen steht noch aus, und es wird weiter verfeinerte experimentelle Szenarien brauchen, um zu klaren Erkenntnissen über den Umgang mit der Ressource Zeit im Entscheidungsverhalten zu kommen.

Wir fanden grundsätzlich unterschiedliche Stopp-Strategien: One-Reason-Decision-Maker reagieren anders auf die Einführung der Zeitachse und der Stabilitätsinformation als More-Reason-Decision-Maker. So erscheint es uns aussichtsreich, bei zukünftigen Experimenten in Design und Auswertung stärker auf diese Distinktion der Strategien (frugal versus absichernd) zu setzen. Diese relativ stabile Präferenz auf die eine oder andere Art die Informationssuche

zu stoppen, wurde schon auch schon von Bröder gefunden (Bröder, 2000a, 2000b) und hier mehrmals repliziert. Die Versuchspersonen wenden die Stoppregel nicht sehr adaptiv in Bezug auf sich verändernde statistische Umwelten an. Vor diesem Hintergrund scheint es ein guter Ausgangspunkt für die weitere Forschung zu sein, wenn man etwas über die grundlegende Einstellung einer Person zur Stoppregel weiss. Zudem erhärtet dieser Befund die Kritik an der Adaptivität der Toolbox. Neben der Beschreibung der einzelnen kognitiven Werkzeuge in den verschiedenen Umwelten fehlt bei der Toolbox-Metapher nach wie vor der Aspekt der Werkzeugselektion (wie dies auch Bröder & Newell, 2008 kritisch bemerken). Unsere Befunde stützen diese Kritik und lassen zudem vermuten, dass dieser Selektionsprozess einerseits durch einen Explorationseffekt in einer ersten statistischen Umwelt, andererseits durch einen Trägheitseffekt (wie ihn auch Rieskamp, 2006 fand) in der nächsten statistischen Umwelt beeinflusst wird. Gigerenzer postuliert in seinem neusten Buch (Gigerenzer 2007), dass die Intuition diese Selektion steuere. Wie diese Hypothese jedoch zu falsifizieren bzw. verifizieren ist, scheint mir zur Zeit noch ein ungelöstes Rätsel. Wir fanden starke Grenzen der Adaptivität der Werkzeugselektion bei der Veränderung der statistischen Umwelt, diese sollten in Bezug auf die Toolbox-Metapher und die verschiedenen statistischen Umwelten in der zukünftigen Forschung noch präziser eruiert werden.

Es wäre jedoch überzogen, die gefundenen Resultate auf die Informationssuche in zeitlich dynamischen Systemen im Allgemeinen zu erweitern, handelt es sich hier doch um eine künstlich dargestellte Zeitdimension. Unterstrichen haben sollten die vorgelegten Studien aber, dass hier eine interessante Forschungsaufgabe besteht, um die Entscheidungsforschung in Experiment und Theoriebildung näher an die realen Bedingungen menschlichen Entscheidens (in der Zeit nämlich) heranzuführen.

Ausblick

Die Metapher der Linse, des Sackmessers oder der Toolbox sind sicherlich hilfreich zur Modellierung der Funktionsweise der menschlichen Psyche und der Verarbeitung von Informationen vor einer Entscheidung. Nichtsdestotrotz zeigen unsere Befunde, dass diese Theorien durch experimentelle Verhaltenstests noch eindringlicher auf ihre Korrektheit überprüft werden sollten, gerade was den zentralen Punkt der Adaptivität angeht. Solche Überprüfungen sollen nicht die Modelle negieren, sondern helfen, ihre Grenzen und Möglichkeiten festzustellen und je nach dem zu generalisieren oder spezifizieren. In diesem Zusammenhang stellen sich noch unzählige offene Forschungsfragen, die es in der zukünftigen Forschung zu beantworten gilt, um den Entscheidungsprozess besser verstehen zu können.

Literatur

- Bergert, F. B., & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and take-the-best model of decision making. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 33, 107-129.
- Bröder, A. (2000a). Assessing the empirical validity of the "Take The Best"-heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1332-1346.
- Bröder, A. (2000b). "Take the best - Ignore the rest" Wann entscheiden Menschen begrenzt rational? Lengerich: Pabst.
- Bröder, A. (2002). Take The Best, Dawes' Rule, and compensatory decision strategies: A regression-based classification method. *Quality & Quantity*, 36, 219-238.
- Bröder, A. (2003). Decision making with the "adaptive toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 611-625.
- Bröder, A. (2005). *Entscheiden mit der "adaptiven Werkzeugkiste". Ein empirisches Forschungsprogramm*. Lengerich: Pabst Science Publishers.
- Bröder, A., & Gaissmaier, W. (2007). Sequential processing of cues in memory-based multi-attribute decisions. *Psychonomic Bulletin and Review*, 14, 895-900.
- Bröder, A., & Newell, B. (2008). Challenging some common beliefs: Empirical work within the adaptive toolbox metaphor. *Judgement and Decision Making*, 3(3), 205-214.
- Bröder, A., & Schiffer, S. (2003). Take The Best versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General*, 132, 277-293.
- Bröder, A., & Schiffer, S. (2006a). Adaptive flexibility and maladaptive routines in selecting fast and frugal decision strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(4), 904-918.
- Bröder, A., & Schiffer, S. (2006b). Stimulus Format and Working Memory in Fast and Frugal Strategy Selection. *Journal of Behavioral Decision Making*, 19, 361-380.
- Brunswik, E. (Ed.). (1952). *The conceptual framework of psychology* (Vol. 10). Chicago: The University of Chicago Press.
- Bussemeyer, J. R., & Townsend, J. T. (1993). Decision Field Theory: A Dynamic-Cognitive Approach to Decision Making in an Uncertain Environment. *Psychological Review*, 100(3), 432-459.
- Christen, S., Hausmann, D., & Läge, D. (2006). Ist ein Gefühl für Informationsgüte lernbar? Ein Lernexperiment zu Validität, Diskriminationsrate und Success von Informationsquellen. *AKZ-Forschungsbericht* 37, Zürich: Angewandte Kognitionspsychologie.
- Cosmides, L., & Tooby, J. (1994). Beyond intuition and instinct blindness: Toward an evolutionary rigorous cognitive science. *Cognition*, 50, 41-77.
- Czerlinski, J., Gigerenzer, G., & Goldstein, D. (1999). How good are simple Heuristics? In G. Gigerenzer, P. Todd & A. R. Group (Eds.), *Simple Heuristics that make us smart*. New York, Oxford: Oxford University Press.
- Dougherty, M., Franco-Watkins, A., & Thomas, R. (2008). Psychological Plausability of the Theory of Probabilistic Mental Models and the Fast and Frugal Heuristics. *Psychological Review*, 115(1), 199-213.
- Dougherty, M. R., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review*, 106, 180-209.
- Enkvist, T., Newell, B., Juslin, P., & Olsson, H. (2006). On the Role of Causal Intervention in Multiple-Cue Judgment: Positive and Negative Effects on Learning. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 32(1), 163-179.
- Gigerenzer, G. (2007). *Bauchentscheidungen: Die Intelligenz des Unbewussten und die Macht der Intuition*. München: Bertelsmann.
- Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The Take The Best heuristic. In P. M. T. G. Gigerenzer, & the ABC Research Group (Ed.), *Simple Heuristics That Make Us Smart* (pp. 75-95). New York: Oxford University Press.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 506-528.
- Gigerenzer, G., & Todd, P. (1999). Fast and Frugal Heuristics- The Adaptive Toolbox. In G. Gigerenzer, P. Todd & A. R. Group (Eds.), *Simple Heuristics that make us smart*. New York, Oxford: Oxford University Press.
- Gigerenzer, G., Todd, P., & Group, A. R. (Eds.). (1999). *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.
- Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In P. M. T. G. Gigerenzer, & the ABC Research Group (Ed.), *Simple heuristics that make us smart* (pp. 3-34). New York: Oxford University Press.

- Glöckner, A. (2006). Automatische Prozesse bei Entscheidungen. Das dominierende Prinzip menschlicher Entscheidungen: Intuition, komplex-rationale Analyse oder Reduktion? (Vol. 9). Hamburg: Verlag Dr. Kovac.
- Glöckner, A., & Betsch, T. (2008). Modelling option and strategy choices with connectionist networks: Towards an integrative model of automatic and controlled decision making. *Judgment and Decision Making*, 3(3), 215-228.
- Goldstein, D., & Gigerenzer, G. (1999). The Recognition Heuristic: How Ignorance Makes Us Smart. In G. Gigerenzer, P. Todd & A. R. Gigerenzer (Eds.), *Simple Heuristics that make us smart* (pp. 37-58). New York, Oxford: Oxford University Press.
- Goldstein, D., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109, 75-90.
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 356-388.
- Hausmann-Thürig, D. (2004). Informationssuche im Entscheidungsprozess: Die Nützlichkeit von Hinweis-Cues und der Anspruch an Urteilssicherheit. Unpublished Promotionsschrift, Universität Zürich, Zürich.
- Hausmann, D., Christen, S., & Läge, D. (2005). Wie viel bezahlen für eine valide Information? Suchkosten als limitierender Faktor der Informationssuche. *AKZ-Forschungsbericht 7*, Zürich: Angewandte Kognitionspsychologie.
- Hausmann, D., & Läge, D. (2005a). Der Wunsch nach Sicherheit: Wie das Anspruchsniveau an Urteilssicherheit in Entscheidungssituationen unter Unsicherheit das Informationssuchverhalten steuert. *AKZ-Forschungsbericht 9*, Zürich: Angewandte Kognitionspsychologie.
- Hausmann, D., & Läge, D. (2005b). Wenn Limiten zum Strategiewechsel führen: Wechsel der Suchstrategie unter situativen Einflüssen. *AKZ-Forschungsbericht 6*, Zürich: Angewandte Kognitionspsychologie.
- Hausmann, D., & Läge, D. (2008). Sequential evidence accumulation in decision making: The individual desired level of confidence can explain the extent of information acquisition. *Judgement and Decision Making*, 3, 229-243.
- Hoffrage, U., Hertwig, R., & Czienskowski, U. (2003). The ecological rationality of Take The Best's building blocks. Zürich: Swiss Federal Institute of Technology.
- Hogarth, R. M., & Karelaia, N. (2006). "Take-the-Best" and other simple strategies: Why and when they work "well" with binary cues. *Theory and Decision*, 61, 205-249.
- Juslin, P., Olsson, H., & Olsson, A. (2003). Exemplar Effects in Categorization and Multiple-Cue Judgment. *Journal of Experimental Psychology: General*, 132(1), 133-156.
- Juslin, P., & Persson, M. (2002). PROBABILITIES from EXEMPLAR (PROBEX): A "lazy" algorithm for probabilistic inference from generic knowledge. *Cognitive Science*, 26, 563-607.
- Karlsson, L., Juslin, P., & Olsson, H. (2008). Exemplar effects in categorization and multiple-cue judgement. *Judgment and Decision Making*, 3, 244-260.
- Kerstholt, J. H., & Raaijmakers, J. G. W. (1997). Decision making in dynamic task environments. In W. R. Crozier & O. Svensson (Eds.), *Decision Making: Cognitive Models and Explanations* (pp. 205-217). London: Routledge.
- Konert, A., Bruziks, N., Gechter, J., & Bröder, A. (2008). Einfluss der Art des Lernens auf Strategien in probabilistischen Multi-Attribut-Inferenzen. Paper presented at the 50. Tagung experimentell arbeitender Psychologen.
- Läge, D., Hausmann, D., & Christen, S. (2005). Wie viel bezahlen für eine valide Information? Suchkosten als limitierender Faktor der Informationssuche. *AKZ-Forschungsbericht 7*, Zürich: Angewandte Kognitionspsychologie.
- Läge, D., Hausmann, D., Christen, S., & Daub, S. (2005). Was macht einen "guten Cue" aus? Strategien der Informationssuche beim heuristischen Entscheiden unter Unsicherheit. *AKZ-Forschungsbericht 5*, Zürich: Angewandte Kognitionspsychologie.
- Lagnado, D. A., Newell, B. R., Kahan, S., & Shanks, D. R. (2006). Insight and Strategy in Multiple-Cue Learning. *Journal of Experimental Psychology: General*, 135(2), 162-183.
- Martignon, L., & Hoffrage, U. (1999). Why does One-Reason Decision Making Work? A case Study in Ecological Rationality. In G. Gigerenzer, P. Todd & A. R. Gigerenzer (Eds.), *Simple Heuristics that make us smart*. New York, Oxford: Oxford University Press.
- Mitchell, T. R., & Beach, L. R. (1990). "...Do I love thee? Let me count ...": Toward an understanding of intuitive and automatic decision making. *Organizational Behavior & Human Decision Processes*, 47, 1-20.
- Newell, B., Rakow, T., Weston, N., & Shanks, D. (2004). Search strategies in decision making: The success of 'success'. *Journal of Behavioral Decision Making*, 17, 117-137.
- Newell, B., & Shanks, D. (2003). Take the best or look at the rest? Factors influencing one-reason decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 53-65.
- Newell, B., Weston, N., & Shanks, D. (2003). Empirical tests of a fast and frugal heuristic: not everyone "takes-the-best". *Organizational Behavior and Human Decision Processes*, 91, 82-96.
- Pachur, T., & Olsson, H. (2008). Strategiewahl in Inferenzaufgaben: Der Einfluss von absolutem und komparativem Feedback. Paper presented at the Beiträge zur 50. Tagung experimentell arbeitender Psychologen.

- Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, 16, 366-387.
- Payne, J. W., & Bettman, J. R. (2000). Preferential Choice and Adaptive Strategy Use. In G. Gigerenzer & R. Selten (Eds.), *Bounded Rationality: The Adaptive Toolbox*. Cambridge, London: MIT Press,.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 534-552.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press.
- Rakow, T., Newell, B., Fayers, K., & Hersby, M. (2005). Evaluating Three Criteria for Establishing Cue-Search Hierarchies in Inferential Judgment. *Journal of Experimental Psychology / Learning, Memory & Cognition*, 31(5), 1088-1104.
- Rieskamp, J. (2006). Perspectives of Probabilistic Inferences: Reinforcement Learning and an Adaptive Network Compared. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(6), 1355-1370.
- Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics and how can we tell? In G. Gigerenzer, P. Todd & A. R. Group (Eds.), *Simple Heuristics That Make Us Smart* (pp. 141-168). New York: Oxford University Press.
- Rieskamp, J., & Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica*, 127, 258-276.
- Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135, 207-236.
- Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (1999). Multi-Alternative Decision Field Theory: A Dynamic Artificial Neural Network Model of Decision-Making. Unpublished manuscript, Indiana University.
- Rottenstreich, Y., & Kivetz, R. (2006). On decision making without likelihood judgment. *Organizational Behavior and Human Decision Processes*, 101, 74-88.
- Schoppek, W., & Putz-Osterloh, W. (2004). Informationsverhalten. In G. Schreyögg & A. Werder (Eds.), *Handwörterbuch Unternehmensführung und Organisation*. Stuttgart: Schäffer-Poeschel.
- Simon, H. A. (1955). A Behavioral Model of Rational Choice. *Quarterly Journal of Economics*, 69, 99-118.
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185, 1124-1131.
- Zurbruggen, S., Christen, S., Hausmann, D., & Läge, D. (2007a). Adaptivität der Informationssuche in statischen und in dynamischen Umwelten. *AKZ-Forschungsbericht 46*, Zürich: Angewandte Kognitionspsychologie.
- Zurbruggen, S., Christen, S., Hausmann, D., & Läge, D. (2007b). Informationssuche beim Entscheiden in zeitlich dynamischen Umwelten. *AKZ-Forschungsbericht 44*, Zürich: Angewandte Kognitionspsychologie.
- Zurbruggen, S., & Läge, D. (2007). Ressource Zeit statt Geld: Ein Vergleich zweier Szenarien zur Entscheidungsforschung. *AKZ-Forschungsbericht 45*, Zürich: Angewandte Kognitionspsychologie.

DIE INFORMATIONSKONSISTENZ

Kapitel 1: Wie zufrieden machen Informationen? Die Bewertung von Informationen in Bezug auf Urteilssicherheit und Urteilszufriedenheit

Einleitung

In der bisherigen Entscheidungsforschung lag der Fokus lange Zeit auf der Entscheidung selbst und prägte die Diskussion um die Rationalität des Menschen. Fehlentscheidungen wurden dabei als irrational oder eben als Fehler (Heuristic and Biases- Programme von Kahnemann und Tversky) gesehen. Mit der Wiederentdeckung von Simons Ansatz der *Bounded Rationality* wurde aus der Irrationalität eine durch Ressourcen begrenzte Rationalität und somit dem Umstand Rechnung getragen, dass Menschen nicht wie Computer alle Informationen optimal verrechnen à la Homo oeconomicus, sondern ihre Ressourcen begrenzt sind durch Gedächtniskapazität oder Zeit. Mit dieser Theorie wechselte der Fokus von der Entscheidung auf den Prozess vor der Entscheidung und auf die Strategien, mit welchen Menschen mit Informationen umgehen, um eine Entscheidung zu treffen. Es etablierte sich unter anderem die ABC-Forschungsgruppe um Gigerenzer, die mit der Toolbox-Metapher adaptive Heuristiken im Informationsverarbeitungsprozess postulierten (Gigerenzer & Goldstein, 1996; Gigerenzer, Hoffrage, & Kleinböling, 1991; Gigerenzer, Todd, & ABC- Research Group, 1999). Als die fundamentalen drei Bausteine dieser Heuristiken gelten die Such-, die Stopp- und die Entscheidungsregel. Sowohl in Simulationsstudien (z.B. Czerlinski, Gigerenzer, & Goldstein, 1999; Hogarth & Karelaia, 2006), als auch in experimentellen Verhaltensforschungen (z.B. Bröder, 2005; Hausmann & Läge, 2005b; Newell, Weston, & Shanks, 2003) wurde seither untersucht, welche und wie viele Informationen gesucht werden, um Entscheidungen zu treffen. Als Gütemass der Informationen wurde unter anderem die Qualität von probabilistischen Cues, z.B. anhand derer Validität oder Diskriminationsrate untersucht. Eine der Heuristiken von Gigerenzer und Kollegen, die *Take The Best* (TTB) Heuristik (Gigerenzer & Goldstein, 1996), postuliert eine Informationssuche gemäss absteigender ökologischer Validität der Cues (Suchregel), bis eine diskriminierende Information gefunden wird und dann soll die Informationssuche gestoppt werden (Stoppregel). Dieses Stoppen wird dann auch als *One Reason Decision Making* (Stoppen nach dem ersten Cue) bezeichnet. Im Gegensatz dazu wird das Fortsetzen der Suche nach einer diskriminierenden Information als *More Reason Decision Making* bezeichnet. Wie es scheint, haben Individuen verhältnismässig stabile Präferenzen für das Entscheiden auf die eine oder die andere Weise (Bröder, 2000a, 2000b; Zurbriggen, Christen, Hausmann, & Läge, 2008a, 2008b; Zurbriggen & Läge, 2007). Bisherige Befunde von Hausmann & Läge (2005a) zeigen eine Schwelle der Konfidenz, d.h. ein individuelles Anspruchsniveau an Urteilssicherheit (Desired Level of Confidence = DLC) bei einer Information, welche die Weitersuche bzw. das Stoppen der Informationssuche bestimmt. Dieses Anspruchsniveau wurde jeweils anhand der Validität der ersten akzeptierten bzw. nichtakzeptierten Information geschätzt. Dabei wurde allerdings bei einer weiteren Informationssuche die Informationskonsistenz von den Informationen nicht in die Schätzung miteinbezogen. Das heisst, dass die Tatsache, ob zwei Informationen sich gegenseitig bestätigen (hohe Informationskonsistenz) oder widersprechen (niedrige Informationskonsistenz) nie variiert wurde und nicht untersucht wurde, inwiefern diese Tatsache das weitere Suchen beeinträchtigt. Mathe-

matische Modelle können zwar berechnen, wie zwei Informationen miteinander integriert werden (z.B. gibt es verschiedene kompensatorische Modelle, wie lineare Regressionen, WADD = weighted additive oder Franklins Rule), doch wie die Personen diese beiden Informationen bewerten und wie diese Bewertung die weitere Informationssuche und das Stoppen bestimmt, ist unserens Wissens bis anhin ungeklärt. Der dritte Baustein der Heuristik, die Entscheidungsregel, ist bei einem *One Reason Decision Making* relativ trivial: die Entscheidung fällt jeweils in die angezeigte Richtung der einen Information. Wenn es nun aber mehrere Informationen sind, dann ist die Entscheidungsregel nicht mehr so trivial. Dieser Problematik wird hier jedoch nicht genauer nachgegangen, da uns hauptsächlich die Informationssuche und das Stoppen interessieren.

Um die Informationssuche und das Stoppen der Informationssuche bei unterschiedlicher Informationskonsistenz in der prädezisionalen Phase genauer erklären zu können, sollen in dieser Untersuchung zwei unabhängige Variablen erhoben werden. Wir nehmen an, dass ein erreichter Grad an Zufriedenheit mit der Informationslage die prädezisionale Phase mitprägt. Diese Zufriedenheit kann sich aus mehreren Komponenten zusammensetzen, von denen wir die Urteilssicherheit V (operationalisiert als Validität) und die Informationskonsistenz (operationalisiert als Grad der Widerspruchsfreiheit von Cues) als zwei der zentralen Faktoren vermuten. Daneben sind andere Faktoren vermutlich auch noch einflussgebend, wie eine Kosten-Nutzen-Analyse bei der Informationssuche (z.B. zeigten Hausmann & Läge (2005b), dass Versuchspersonen die Stopppregel auf eine nicht mathematisch korrekte Weise ändern, wenn sich die Informationskosten verändern). Die Zufriedenheit definieren wir als eine kontinuierliche Konstruktvariable, genannt „Urteilszufriedenheit“, wobei wir einen Sprung der Linearität beim Erreichen des DLCs vermuten. Das Ziel dieser Untersuchung ist, den Einfluss der Urteilszufriedenheit auf die zwei Bausteine Such- und Stopppregel explorativ zu ergründen. Einerseits soll dabei der Zusammenhang zwischen subjektiver Urteilssicherheit und Urteilszufriedenheit bei einer Information gemessen werden und andererseits der Einfluss der Informationskonsistenz auf diese Faktoren bei zwei Informationen erfasst werden. Der erste Teil des Experimentes (Phase 1) dient zur Erhebung der Basisbedingung in Bezug auf die Bewertung und das Informations- und Stoppverhalten bei einer Information. Im zweiten Teil des Experimentes (Phase 2) werden dann zwei Informationen, mit unterschiedlicher Informationskonsistenz, in Bezug auf die Urteilssicherheit und Urteilszufriedenheit bewertet. Anhand dieser Methodik kann die Bewertung und deren Einfluss auf den prädezisionale Verhaltensprozess bei einer Information und bei zwei Informationen miteinander verglichen werden.

Methoden

Stichprobe

30 Personen, darunter 19 Psychologie-Studierende der Universität Zürich nahmen an diesem Experiment teil. Etwas mehr als die Hälfte der TeilnehmerInnen waren Frauen (16 Frauen, 14 Männer). Das Durchschnittsalter betrug 27.3 Jahre (Range 16–57, SD = 9.8). Die Psychologie-Studierenden müssen im Rahmen ihrer Ausbildung an der Universität Zürich mindestens fünf Stunden an Experimenten teilnehmen, an welchen ist ihnen dabei frei gestellt. Die Nicht-Studierenden nahmen freiwillig teil. Unter dem besten Drittel wurde zu Motivationszwecken zweimal eine Gewinnsumme von CHF 100.- verlost.

Vorgehen

Das Experiment wurde in Microsoft Visual Basic 6.0 programmiert und vollständig auf einem portablen PC mit Windows XP-Betriebssystem durchgeführt. Das Experiment war anhand eines Information Boards aufgebaut (vgl. Abbildung 1) und entsprach damit der grundsätzlichen Logik des Mouselabs von Payne et al. (Payne, Bettman, & Johnson, 1993).

Pferde
Pferderennen Nr.
Die folgenden vier Pferde stehen am Start:

Tipps von Experten
V-Rate = Mittlere Anzahl korrekter Vorhersagen pro 100 Pferderennen

Experte	V-Rate	Tipp
Koch	87	Epimetheus
Werner	28	CHF 30.-
Steiner	82	CHF 30.-
Huber	72	CHF 30.-
Weber	84	CHF 30.-
Richter	69	CHF 30.-

Sicherheit
Wie sicher bist Du, dass Epimetheus gewinnt?
gewinnt sicher 100%
90%
80%
70%
60%
50%
40%
30%
25%
20%
10%
0%
Chance wie jedes andere Pferd

Zufriedenheit
Angenommen, Du müsstest Dich nun entscheiden, wie zufrieden bist Du mit der aktuellen Informationslage?

Was möchtest Du nun tun?

Pferderennen
Gewonnen hat:
Dein Tip wer:

Entscheidung
Wenn Du jetzt auf ein Pferd setzen müsstest, für welches würdest Du Dich entscheiden?

Kontostand [CHF]
Aktueller Kontostand:
Kosten für Experten-Tipps: -
Wett-Gewinn: +
Neuer Kontostand: =

Abbildung 1: Informationboard Phase 1

Die Versuchspersonen lösten anhand des Information Boards insgesamt 70 Entscheidungsaufgaben. Das Szenario basierte auf dem fiktiven Pferderennszenario von Hausmann & Läge (2005a) mit vier Optionen und sechs Informationsquellen. Die Versuchsperson musste auf eines von vier Pferden setzen, wobei sechs verschiedene Experten nach ihrer Meinung befragt werden konnten. Die vier Pferdenamen wurden jeweils zufällig aus einem Pool von insgesamt 288 Namen (zusammengestellt aus griechischen, römischen, ägyptischen, keltischen und germanischen Göttern oder mythologischen Gestalten) gezogen. Es gab insgesamt 50 Experten, welche alle mit den häufigsten Nachnamen der Schweiz und Deutschlands bezeichnet waren. Jeder Experte hatte eine bestimmte Vorhersagerate (V-Rate), d.h. eine Zuverlässigkeit, welche in ganzen Zahlen auf dem Information Board ersichtlich war. Die Zuverlässigkeit stand für die mittlere Anzahl korrekter Vorhersagen, die dieser Experte pro 100 Pferderennen bisher gemacht hatte. Diese Vorhersagerate ist die experimentelle Implementierung des Konzeptes der ökologischen Validität der Cues von Gigerenzer et al. (1991). Jedoch haben die personifizierten Cues den Vorteil, dass sie als voneinander unabhängig betrachtet werden können und man auch die Diskriminationsrate, im inhaltlichen Sinne einer Erreichbarkeit, im experimentellen Setting frei variieren könnte. In diesem Experiment wurde die Diskriminationsrate jedoch bei der maximalen Ausprägung von 1.00 konstant gehalten, d.h. wenn immer ein Experte nach seiner Meinung befragt wurde, gab dieser auch einen Tipp ab. Die Experten konnten von den Versuchspersonen frei befragt werden. Die Reihenfolge (Bildschirmposition) der Experten wurde während der Laufzeit des Experiments zufällig be-

stimmt. Um eine Auswertung anhand des Erstwertes zu garantieren, musste immer mindestens ein Experte befragt werden. Sobald ein Name eines Experten angeklickt wurde, erschien dessen Tipp, also der Hinweis auf ein Siegerpferd. Dieser Tipp sagte mit der Wahrscheinlichkeit der Vorhersagerate den richtigen Sieger voraus. Nach jedem Tipp musste eine hypothetische Entscheidung getroffen werden, die Sicherheit über die zu erwartenden Gewinnchancen und die Zufriedenheit mit der aktuellen Informationslage angegeben werden. Das Sicherheitsrating musste auf einer vertikalen Skala von 0% bis 100% eingestellt werden, das Zufriedenheitsrating auf einer horizontalen 17 stufigen Skala von minimaler bis maximaler Zufriedenheit (illustriert mit einem Smiley). Danach konnten entweder weitere Experten befragt werden oder die Informationssuche beendet und auf das Pferd getippt werden. Nach gelaufenem Rennen wurde der Sieger angezeigt und eine Rückmeldung gegeben, ob man mit der Entscheidung richtig oder falsch lag, worauf dann das nächste Rennen folgte. Während dem ganzen Spiel war der aktuelle Kontostand ersichtlich, der jeweils einsetzbare Betrag von 1000.- und die abgezogenen Informationssuchkosten. Das Experiment bestand aus zwei Phasen, welche im Folgenden beschrieben werden.

Phase 1: Bewertung einer Information und Schätzung des DLCs

Auf eine erste Trainingsrunde folgten 30 Aufgaben, welche einerseits zur Schätzung der Höhe des Anspruchsniveaus an individueller Urteilssicherheit (DLC) dienten und andererseits zur Messung der Urteilssicherheit und Urteilszufriedenheit von einer Cue-Information, mit variierender Validität zwischen 0.31 bis 0.96 (insgesamt 30 Validitätswerte).

Urteilssicherheits- und Urteilszufriedenheitsratings

Da die Versuchspersonen den ersten Cue frei wählen konnten, wurden die Validitäten der sechs Cues auf dem Information Board so zusammengestellt, dass jeder der 30 Validitätswerte einmal der beste war. Die Validitäten der übrigen fünf Cues wurden während dem Experiment zufällig ausgewählt und dargestellt. Eine Information kostete 30.-, welche nach dem Anklicken vom Budget abgezogen wurde. Nach dem Erhalt des ersten Hinweises mussten die Versuchsteilnehmer auf ein Pferd setzen, die subjektive Gewinnchance einschätzen und die Zufriedenheit mit der Informationslage. Dies ergab somit für die 30 Validitätswerte sowohl einen subjektiven Urteilssicherheitswert, wie auch ein Zufriedenheitsrating. Danach konnten die Versuchspersonen entweder auf das Pferd setzen oder weitere Informationen suchen bis sie auf ein Pferd setzten. Diese weiteren Ratings wurden jedoch nicht weiter ausgewertet. Die Coverstory für Phase 1 lautete folgendermassen:

Mit Pferdewetten betreibst du einen lukrativen Nebenjob. Wenn das Pferd, auf welches du gesetzt hast, das Rennen gewinnt, wächst dein Spielkonto um 300 Franken an. Wenn du auf das falsche Pferd gesetzt hast, gehst du leer aus.

Solange du nichts über die 4 startenden Pferde weißt, hat jedes Pferd grundsätzlich dieselbe Wahrscheinlichkeit, das Rennen zu gewinnen. Indem du jeweils vor den Rennen mit verschiedenen Experten, also z.B. Pferdebesitzern, Jockeys, Rennveranstaltern etc. redest, erhältst du wertvolle Insiderinformationen über die Gewinnwahrscheinlichkeit der Pferde.

Wenn du die Experten nach ihrer Meinung fragst, geben sie dir an, welches der Pferde sie selber als Sieger des aktuellen Rennens sehen. Insgesamt gibt es 50 Experten, die sich allerdings in ihrem Expertenwissen voneinander unterscheiden. Von jeder Person weisst du jedoch, wie viele korrekte Vorhersagen sie in den letzten 100 Rennen getroffen hat. Diese

Zahl schwankt zwischen 0 bis 100 korrekter Vorhersagen. Bei jedem Rennen sind von den 50 Experten, jeweils sechs anwesend.

Das Einholen von Insiderinformationen ist mit gewissen Kosten verbunden. Das Befragen einer einzelnen Person kostet dich jeweils 30 Franken und wird dir direkt vom Spielkonto abgebogen. Dein Startkapital zu Spielbeginn beträgt 1'000 Franken. Für jedes einzelne Rennen kannst du bis zu sechs Experten befragen, mindestens eine Person musst du auf jeden Fall befragen!

Nachdem du jeweils einen Experten befragt hast, möchten wir gerne von Dir wissen, wie sicher du dir mit dessen Tipp bist und wie zufrieden dich diese Information macht. Bitte fülle dies sorgfältig aus.

Es folgen 30 Pferdewetten, versuche dabei einen möglichst hohen Gewinn zu erzielen! Die erste Aufgabe ist zum kennen lernen des Settings da und zählt noch nicht.

Schätzung des Desired Level of Confidence (DLC)

Hausmann (2004) entwickelte ein Schätzverfahren, welches anhand der akzeptierten und der zurückgewiesenen (d.h. weiteres Einholen von Expertenmeinungen) Erstwerte der Validitäten ein hypothetisches individuelles Anspruchsniveau an Urteilssicherheit (DLC) berechnet. „Ein Computeralgorithmus berechnete aus den akzeptierten und zurückgewiesenen Erstwerten den quadrierten Fehlerwert für jedes potenzielle ganzzahlige Anspruchsniveau zwischen .00 und 1.00. Als Fehler wurde gezählt, wenn höhere Wahrscheinlichkeitswerte (Erstwerte) als das hypothetische Niveau fälschlicherweise ignoriert (die Vpn hat weitere Hinweise aufgedeckt) wurden oder Erstwerte unterhalb der Sicherheitsschwelle fälschlicherweise akzeptiert (die Vpn hat keine weiteren Hinweise aufgedeckt) wurden. Bei einem Verstoß gegen das Abbruchkriterium in die eine oder andere Richtung definierte sich die Höhe des Fehlerwertes als Differenzwert zwischen dem hypothetischen Niveau und dem präsentierten Erstwert. Als hypothetisches individuelles DLC wurde der Prozentwert mit der minimalen Fehlersumme angenommen“ (S.96, Hausmann, 2004). Da aber dieses bisherige Schätzverfahren des DLCs anhand der summierten Abweichungsquadrate (SAQ) nicht robust auf Ausreisser ist, wurde dieses Schätzverfahren leicht adaptiert. Neu wurden die einzelnen Fehler gezählt und aufsummiert und erst bei zwei oder mehreren lokalen Minima dieser Fehlersumme wurde die Methode der summierten Abweichungsquadrate verwendet, um das individuelle Desired Level of Confidence zu schätzen. Eine weitere Änderung im Experiment bestand darin, dass die Validitäten aller sechs Experten angegeben wurden, und nicht wie zuvor immer nur ein Experte ersichtlich war und die weiteren Experten sequentiell aufgedeckt werden mussten. Daher basiert die Einschätzung des DLC nicht auf einem vorgegebenen Erstwert, sondern auf dem Wert, den die Person als erstes aufdeckt.

Phase 2: Bewertung von zwei Informationen

In Phase 2 gaben immer zwei Experten gleichzeitig ihre Tipps ab. Die Coverstory lautete:

Nun spielst Du nochmals dasselbe Spiel wie zuvor. Du triffst jedoch immer auf zwei Experten zusammen. Um keinen der beiden zu beleidigen, befragst du sie jeweils zu zweit. Den Tipp erhältst Du jedoch von beiden getrennt. Das Befragen der beiden Personen kostet Dich jeweils 50 Franken und wird Dir direkt vom Spielkonto abgebogen, der Gewinn bei einem richtigen Tipp beträgt 300 Franken. Ansonsten ändert sich nichts! Es folgen nochmals 40 Aufgaben, versuche weiterhin möglichst viel Gewinn zu machen!

Die erste Aufgabe ist wiederum zum kennen lernen des Settings und zählt noch nicht.

Pferde

Pferderennen Nr.

Die folgenden vier Pferde stehen am Start:

Bes	Epona
Uto	Polyhymnia

Tipps von Experten

V-Rate = Mittlere Anzahl korrekter Vorhersagen pro 100 Pferderennen

Experte	V-Rate	Tipps
Schneider	66	CHF 50.-
Roth	53	
Jäger	55	CHF 50.-
Meier	38	
Weber	84	Epona
Gerber	54	Polyhymnia

Entscheidung

Wenn Du jetzt auf ein Pferd setzen müsstest, für welches würdest Du Dich entscheiden?

Bes	Epona
Uto	Polyhymnia

Sicherheit

Wie sicher bist Du, dass Epona gewinnt?

gewinnt sicher

100% 90% 80% 70% 60% 50% 40% 30% 25% 20% 10% 0%

Chance wie jedes andere Pferd

84%

OK

Zufriedenheit

Angenommen, Du müsstest Dich nun entscheiden, wie zufrieden bist Du mit der aktuellen Informationslage?

min. max.

OK

Was möchtest Du nun tun?

Weitere Expertenmeinung einholen? Auf Epona tippen?

Pferderennen

Gewonnen hat:

Dein Tip war:

Kontostand [CHF]

Aktueller Kontostand:	1'000
Kosten für Experten-Tipps:	- 50
Welt-Gewinn:	+ ?
Neuer Kontostand:	= ?

Weiter

Abbildung 2: Informationboard Phase 2

Nach dem Trainings-Durchgang absolvierten die Versuchspersonen 40 Runden. Pro Runde standen wiederum sechs Experten zur Verfügung, diesmal allerdings in der Form dreier Expertenpaare (d.h. als drei Paare zu je zwei Experten). Die Experten konnten diesmal nur als Paar befragt werden. Sie gaben entweder widersprechende (d.h. beide Experten waren unterschiedlicher Meinung) oder bestätigende (d.h. beide Experten waren gleicher Meinung) Tipps ab. Die Validitäten der beiden Experten wurden so festgelegt, dass ihre A-posteriori-Validität (siehe Anhang Exkurs A-posteriori-Validität versus A-priori-Validität) jeweils dem Validitätswert eines Experten aus Phase 1 entsprach. Ob ein Expertenpaar eine hohe oder niedrige Informationskonsistenz aufwies, war ebenfalls bereits vor der Experimentaldurchführung festgelegt worden. Die Reihenfolge (Bildschirmposition) der beiden Experten eines Expertenpaares wurde wie in Phase 1 erst während der Laufzeit des Experiments zufällig bestimmt. Auch die Tipps der Experten wurden in Echtzeit, d.h. erst während des Experiments, gemäss folgender Prozedur generiert:

Zunächst wurde ein zufälliger Validitätswert erzeugt. Dieser wurde mit der A-posteriori-Validität der beiden Experten verglichen. War der Zufallswert kleiner oder gleich dieser A-posteriori-Validität (d.h. der Wahrscheinlichkeit mit der die beiden Experten auf das siegreiche Pferd tippen), so gab der erste Experte einen korrekten Tipp ab. War der Zufallswert grösser als die Expertenvalidität, so gab der erste Experte einen falschen Tipp ab (d.h. einen Tipp auf eines der drei nicht-siegreichen Pferde). Dabei war die Wahrscheinlichkeit, ausgelost zu werden, für jedes der drei Pferde gleich hoch (genau ein Drittel der Gegenwahrscheinlichkeit der A-posteriori-Validität). Bei einem Expertenpaar mit hoher Informationskonsistenz musste der so generierte Tipp anschliessend noch dem zweiten Experten zugewiesen werden. Bei einem Expertenpaar mit niedriger Informationskonsistenz, wurde hingegen solange ein neuer Tipp generiert, bis sich der neue Tipp vom ersten Tipp unterschied. Erst danach konnte dieser dem zweiten Experten zugewiesen werden.

Das restliche experimentelle Setting der Phase 2 entsprach der Phase 1: nach dem Aufdecken zweier Tipps war ein Urteilssicherheits- und ein Urteilszufriedenheitsrating anzugeben. Danach konnten entweder noch weitere Expertenpaare befragt oder auf ein Pferd gesetzt werden. Dies ergab für 18 Validitätswerte aus Phase 1 korrespondierende subjektive Urteilssicherheits- und Urteilszufriedenheitsrating von jeweils einer hohen und einer tiefen Informationskonsistenz.

Fragestellung

Um den Zusammenhang zwischen der Urteilszufriedenheit und den verschiedenen Komponenten wie Validität, Urteilssicherheit und Informationskonsistenz einerseits und deren Einfluss auf das Such- und Stoppverhalten in der prädezisionalen Phase andererseits zu untersuchen, stellen wir folgende Fragestellungen in den Vordergrund:

1. Bewertung der Information

Es sollte ein Zusammenhang zwischen der subjektiven Urteilssicherheit und der Validität des ersten konsultierten Cues in Phase 1 gegeben sein. Sobald die Validität einer Information dem individuellen Anspruchsniveau an Urteilssicherheit entspricht, postulieren wir einen Zusatznutzen und somit eine Abweichung der Urteilszufriedenheit von der Urteilssicherheit. Die Urteilszufriedenheit sollte dann, sobald das gewünschte Anspruchsniveau an Urteilssicherheit erreicht ist, ansteigen. Dies würde bedeuten, dass die Personen die Urteilszufriedenheit vor dem Erreichen des individuellen DLCs tiefer einschätzen als die Urteilssicherheit und nach dem Erreichen ihres DLCs die Urteilszufriedenheit höher einstufen als die Urteilssicherheit. Auf Mittelwertebene sollte sich daher ein signifikanter Unterschied zwischen Urteilssicherheit und Urteilszufriedenheit unter- und oberhalb der individuellen DLCs finden lassen. Zudem sollte die Urteilszufriedenheit – bei gleicher Urteilssicherheit – höher bewertet werden, wenn die Informationssuche nach dem ersten Cue beendet wird als wenn weitergesucht wird.

In Phase 2 nehmen wir an, dass die Versuchspersonen zwei Informationen gemäss dem Bayes-Theorem verrechnen und somit die Urteilssicherheit der A-posteriori-Validität (siehe Anhang Exkurs A-posteriori-Validität versus A-priori-Validität S.38ff) des Expertenpaares entspricht. Bei der Urteilszufriedenheit gehen wir von einem zusätzlichen Einfluss der Informationskonsistenz aus. Die Informationskonsistenz ist niedrig, wenn zwei Informationen einander widersprechen, und sie ist hoch, wenn die beiden Informationen einander bestätigen. Bei niedriger Informationskonsistenz sollte die Urteilszufriedenheit tiefer sein als bei hoher Informationskonsistenz. Damit das Konzept der Urteilszufriedenheit einen zusätzlichen Erklärungswert hat, sollte die Informationskonsistenz keinen Einfluss auf die Urteilssicherheit haben, sondern nur auf die Urteilszufriedenheit.

2. Informationssuche

Gemäss der Urteilsheuristik von *Take The Best* (TTB) sollte in der Phase 1 immer der valideste Experte angefragt werden. In der Phase 2 sollte, TTB als Normfall annehmend, das Expertenpaar mit dem validesten Experten konsultiert werden. Integrieren die Versuchspersonen die beiden Validität des Expertenpaares gemäss der berechneten A-priori-Validität, dann sollten sie ebenfalls immer das Paar mit dem validesten Experten konsultieren (da die berechnete A-priori-Validität eines Paares immer der Validität des besseren Experten entspricht, siehe

Exkurs). Integrieren die Personen jedoch die Validitäten der beiden Experten nicht gemäss der berechneten A-priori-Validität, dann sollte auch ab und zu ein anderes Expertenpaar konsultiert werden.

3. Stoppen der Informationssuche

In Phase 1 sollte gemäss TTB nach der ersten Information (da die Information in dem Experiment immer diskriminieren und die Experten immer einen Tipp abgeben) die Informationssuche gestoppt werden. Wir nehmen jedoch ein individuelles DLC an, gemäss dem die Informationssuche beim Erreichen des Anspruchsniveaus an Urteilssicherheit gestoppt werden sollte.

Wird nun aber weitergesucht, weil das Anspruchsniveau noch nicht erreicht ist, dann sollte in Phase 2 die Informationskonsistenz die Weitersuche bzw. das Stoppen vorhersagen: bei niedriger Informationskonsistenz sollte eher weitergesucht werden als bei hoher Informationskonsistenz.

Resultate

Schätzung des Desired Level of Confidence (DLC)

Der Mittelwert aller DLCs liegt bei .55, die Standardabweichung bei .19. Von den 24 Versuchspersonen haben sieben ein geschätztes $DLC \leq .40$. Das bedeutet, dass diese Personen ein ziemlich konsequentes *One Reason Decision Making* betreiben, also nach einer Information die Informationssuche stoppen und sich entscheiden. Bei zwei Personen wird das individuelle $DLC \geq .80$ eingeschätzt, sie können als *More Reason Decision Makers* bezeichnet werden. Diese Personen entscheiden sich lieber anhand mehrerer Informationen, wenn die Validitäten gering sind. In Tabelle 3 im Anhang sind die individuellen DLCs dargestellt.

Resultate Phase 1

Bewertung einer Information

Um einen Zusammenhang zwischen der Validität, der subjektiven Urteilssicherheit und der Urteilszufriedenheit zu eruieren, wurden einerseits die Korrelationen gerechnet und andererseits die Mittelwerte der Bewertungen miteinander verglichen.

Bei den individuellen Korrelationen zwischen den drei Variablen Validität, subjektive Urteilsicherheit und Urteilszufriedenheit (siehe Tabelle 3 in Anhang) wird ersichtlich, dass vier Personen unter Validität und Urteilssicherheit etwa gänzlich anderes verstanden (Korrelationen kleiner als .35) und zwei Personen die Urteilszufriedenheit ebenfalls unzusammenhängend mit der Urteilssicherheit bzw. Validität bewerteten. Diese sechs Versuchspersonen haben vermutlich die zu untersuchenden Variablen anders verstanden als von uns intendiert und werden deshalb aus den folgenden Auswertungen ausgeschlossen. Über alle 720 Trials gibt es bei der Bewertung nach einer Information signifikante Korrelationen zwischen der Validität und der subjektiven Urteilssicherheit von $r=.79$, zwischen der subjektiven Urteilsicherheit und der Urteilszufriedenheit von $r=.76$ und zwischen der Validität und der Urteilszufriedenheit von $r=.75$. Die Korrelationen sind alle auf dem Niveau von .01 (einseitig) signifikant.

Bei den inferenzstatistischen Auswertungen (t-Tests für abhängige Stichproben) wurde eine zweiseitige Hypothesenprüfung als Grundlage für das Akzeptieren eines Mittelwertsunterschiedes und eine einseitige Hypothesenprüfung als Grundlage für das Akzeptieren eines Mittelwertunterschiedes in die vorhergesagte Richtung (immer mit $\alpha = .05$) gewählt. Zwischen den Validitäten und den subjektiven Urteilssicherheiten zeigte sich über die 30 Durchgänge ein signifikanter Unterschied: die Versuchspersonen waren unsicherer in ihrem Urteil ($M = 56.6$; $SD = 20.2$) als die Validität des ersten Cues ($M = 66.0$; $SD = 18.1$) implizieren würde ($t(719) = 20.0$; $p(\text{zweiseitig}) < .001$; $d = 0.5$). Sowohl wenn die Validität des ersten Cues über dem individuellen DLC ($M = 73.6$, $SD = 14.8$) lag, war die Urteilssicherheit ($M = 63.5$, $SD = 18.3$) signifikant niedriger ($t(468) = 16.7$; $p(\text{einseitig}) < .001$; $d = 0.61$) als auch wenn die Validität unter dem individuellen DLC lag (Validität $M = 51.9$, $SD = 15.0$; Urteilssicherheit $M = 43.7$, $SD = 17.1$; $t(250) = 11.1$; $p(\text{einseitig}) < .001$; $d = 0.51$). Das bedeutet, dass die Versuchspersonen die Validität einer Information sowohl über ihrem persönlichen DLC als auch unter ihrem persönlichen DLC um etwa den gleichen Betrag unterschätzten. Die Effektstärke bleibt dabei bei allen Vergleichen gemäss Bortz & Döring (1995/2002) im mittleren Bereich.

Um die unterschiedliche Skalenausnutzung bei der Bewertung der Urteilssicherheit und der Urteilszufriedenheit auszuschliessen, werden die Daten zur Testung der Vorhersagen betreffend der Unter- bzw. Überschätzung in Bezug auf das DLC z-transformiert. Nimmt man die Fälle, bei welchen die Urteilssicherheit des ersten konsultierten Cues unter dem individuellen DLC liegen, dann wurde die Urteilszufriedenheit im Mittel höher eingestuft ($M = -.42$; $SD = .90$) als die Urteilssicherheit ($M = -.51$; $SD = .82$; $t(368) = -2.5$; $p(\text{einseitig}) < .01$, $d = 0.10$). Hingegen zeigt sich bei der Bewertung der Cues über den individuellen DLCs, dass die Urteilszufriedenheit ($M = .53$; $SD = .98$) eindeutig niedriger eingeschätzt wurde als die Urteilssicherheit ($M = .73$; $SD = .76$, $t(350) = 6.0$; $p(\text{einseitig}) < .001$, $d = -0.22$). Diese Unterschiede zwischen der subjektiven Urteilssicherheit und der Urteilszufriedenheit über und unter dem individuellen DLC werden zwar signifikant, aber sie zeigen genau in die entgegengesetzte Richtung, wie in den Hypothesen formuliert wurde und bewegen sich in einem Bereich von sehr kleinen Effektstärken.

In den Hypothesen spezifizierten wir, dass beim Erreichen des individuellen Anspruchsniveaus an Urteilssicherheit die Urteilszufriedenheit sprunghaft ansteigt. Daher sollte insgesamt die Urteilszufriedenheit – bei gleicher Urteilssicherheit – höher bewertet werden, wenn die Informationssuche nach dem ersten Cue beendet wird als wenn weitergesucht wird. Diese Hypothese kann bestätigt werden, denn die Versuchspersonen, welche nach einer Information weitersuchten, waren weniger zufrieden mit der aktuellen Informationslage ($M = 9.6$, $SD = 2.0$) als diejenigen, welche nach einer Information die Suche beendeten ($M = 10.3$, $SD = 2.0$). Dieses Resultat basiert auf 109 Datenpaaren und ist mit $p(\text{einseitig}) < .05$ knapp signifikant ($t(108) = -1.7$, $d = -0.35$).

Informationssuche und Stoppen

In der ersten Phase, in welcher nur eine Information gesucht wurde, wurde in 93.8 % aller Fälle (720) der valideste Cue als erstes angeschaut und somit überzufällig häufig gemäss der Suchreihenfolge von TTB diese erste Information konsultiert. Sechs Personen haben konsequent immer den validesten Cue konsultiert, die restlichen 6.2% Fälle (also das Aufdecken

eines weniger validen Cues) scheinen eher regelmässige Abweichungen der übrigen 18 Versuchspersonen zu sein.

Bei einer vorgegebenen sequentiellen Informationssuche stoppten die Versuchspersonen im Durchschnitt die Informationssuche in 65% aller Fälle nach einer Information und trafen eine Entscheidung gemäss einem *One Reason Decision Making* (siehe Tabelle 1). Das individuelle Anspruchsniveau erklärt nun zusätzlich noch diejenigen Fälle, in welchen wegen zu tiefer Urteilssicherheit weitergesucht wurde (MRDM und $US_1 < DLC$). Somit erklärt das Modell des individuellen Anspruchsniveaus 76% des Stoppverhaltens, was signifikant mehr ist als TTB mit den 65% erklärtem Stoppverhalten (Chi-Quadrat Test $p < .001$).

Tabelle 1: Häufigkeiten des Stoppverhaltens in Phase 1 ($n=720$), wobei die Urteilssicherheit (US_1) des ersten Cues jeweils über bzw. unter dem individuellen DLC lag (ORDM= One Reason Decision Making, MRDM= More Reason Decision Making).

	ORDM	MRDM	Summe
$US_1 \geq DLC$	323	28	351
$US_1 < DLC$	145	224	369
Summe	468	252	720

Resultate Phase 2

Bewertung von zwei Informationen

Die Korrelationen und Mittelwerte sind, wenn man den Einfluss der Informationskonsistenz auf die Bewertung der Information betrachten möchte, je nach Informationskonsistenz getrennt auszuwerten. Bei der Bewertung von zwei widersprüchlichen Informationen (niedrige Informationskonsistenz, $N=723$) signifikante Korrelationen zwischen der A-posteriori-Validität und der subjektiven Urteilssicherheit von $r=.73$, zwischen der subjektiven Urteilssicherheit und der Urteilszufriedenheit von $r=.81$ und zwischen der A-posteriori-Validität und der Urteilszufriedenheit von $r=.66$. Bei der Bewertung von zwei sich bestätigenden Informationen (hohe Informationskonsistenz, $N=237$) betragen die Korrelation zwischen der A-posteriori-Validität und der subjektiven Urteilssicherheit von $r=.50$, zwischen der subjektiven Urteilssicherheit und der Urteilszufriedenheit von $r=.81$ und zwischen der A-posteriori-Validität und der Urteilszufriedenheit von $r=.47$. Die Korrelationen sind alle auf dem Niveau von .01 signifikant.

Zwischen den A-posteriori-Validitäten und den subjektiven Urteilssicherheiten zeigt sich über die 40 Durchgänge ein signifikanter Unterschied bei den widersprüchlichen Informationen und bei den bestätigenden Informationen. Sowohl die Validitäten der widersprüchlichen Informationen ($M=69.1$, $SD=17.1$) wurden unterschätzt ($M=62.4$, $SD=19.3$; $t(722)=13.2$, p (einseitig) $< .001$, $d=0.37$), wie auch die Validitäten ($M=76.1$, $SD=15.9$) der bestätigenden Informationen ($M=67.2$, $SD=17.5$, $t(236)=8.1$, p (einseitig) $< .001$, $d=0.53$).

Da aufgrund der vorherigen Auswertungen die Voraussetzung einer Gleichheit von Validität und Urteilssicherheit nicht angenommen werden kann, müssen wiederum die Urteilssicherheiten statt der Validitäten verwendet werden. Für den Vergleich von subjektiver Urteilssicherheit und Urteilszufriedenheit bei bestätigenden und widersprechenden Informationen müssen intraindividuell Wertepaare ausgewählt werden, welche in der Urteilssicherheit gleich sind: das ergibt insgesamt 139 Datenwertpaare. Um auch hier wieder unterschiedliche Nutzung der Skalen auszuschliessen, wurden die alle Werte vor der Auswahl der Datenwertpaare z-

transformiert (und wenn mehrer Werte zur Urteilszufriedenheit bei einer Urteilssicherheit vorlagen wurden diese gemittelt). Wir finden einen signifikanten Einfluss von der Informationskonsistenz auf die Urteilszufriedenheit ($t(138) = -4.7; p < .001, d = 0.31$). Die Personen waren eindeutig zufriedener beim Erhalt von zwei bestätigenden Informationen ($M = .29; SD = .71$) als nach widersprüchlichen Informationen ($M = .04, SD = .91$). Dieser Befund bestätigt unsere Hypothese des Einflusses der Informationskonsistenz auf die Urteilszufriedenheit.

In den Hypothesen spezifizierten wir zudem, dass die Urteilszufriedenheit – bei gleicher Urteilssicherheit und bei derselben Person – höher bewertet werden sollte, wenn die Informationssuche nach dem ersten Cue beendet wurde als wenn weitergesucht wurde. Dieser Einfluss der Urteilszufriedenheit auf das Stoppen kann ebenfalls bestätigt werden: bei den 33 Datenwertpaaren zeigt sich, dass die Personen intraindividuell zufriedener sind wenn sie die Informationssuche stoppen ($M = -0.3, SD = .76$), als wenn sie weitersuchen ($M = -0.91, SD = .74; t(32) = 5.5, p < .001, d = 0.81$). Dieser letztere Befund kann als grosser Effekt interpretiert werden.

Informationssuche und Stoppen

Als in der Phase 2 immer zwei Informationen vorlagen, konsultierten die Versuchspersonen nur noch in 78 % aller Fälle (960) das Expertenpaar gemäss der A-priori-Validität, d.h. sie befragten in 747 Fällen von allen drei zur Verfügung stehenden Informationspaaren dasjenige Expertenpaar mit dem besten Experten. Diese Suche gemäss TTB ist somit signifikant seltener als in der Vergleichsbedingung von einer Information (im Chi-Quadrat Test ein signifikanter Unterschied mit $p < .001$).

In der Phase 2 wurde die Suche in 786 Fällen (siehe Tabelle 2) nach dem ersten konsultierten Expertenpaar gestoppt. Das bedeutet, dass in 82% aller Fälle die Informationssuche nach den beiden vorgegebenen Informationen abgebrochen wurde. Diese Häufigkeit des Stoppens unterscheidet sich signifikant von der Stophhäufigkeit in der Phase 1 (Chi-Quadrat Test mit $p < .001$). Wenn die Suche gestoppt wurde, dann war in der Hälfte dieser Fälle die Information bestätigend und die Urteilssicherheit über dem DLC. Es wurde aber auch gestoppt, wenn die Urteilssicherheit unter dem DLC lag, aber die Informationen einander bestätigten (in 20% des Stoppverhaltens). In der gleichen Anzahlhäufigkeit wurde aber in diesen Fällen von tiefer Urteilssicherheit und niedriger Informationskonsistenz auch weitergesucht. Von den insgesamt 18% Fällen von Weitersuche war dies fast immer bei niedriger Urteilssicherheit und niedriger Informationskonsistenz zu finden. Fast nie, d.h. nur in 0.4% aller Fälle wurde nach hoher Informationskonsistenz weitergesucht.

Tabelle 2: Häufigkeiten des Stoppverhaltens in Phase 2 ($n = 960$), wobei die subjektive Urteilssicherheit des ersten Cuepaares (US_2) jeweils über bzw. unter dem individuellen DLC lag (Stopp = keine weitere Informationssuche, Weiter = kein Stoppen der Informationssuche nach dem ersten Cuepaar; IK = Informationskonsistenz).

	Stopp	Weiter	Summe
$US_2 \geq \text{DLC} - \text{Niedrige IK}$	393	28	421
$US_2 \geq \text{DLC} - \text{Hohe IK}$	157	2	159
$US_2 < \text{DLC} - \text{Niedrige IK}$	160	142	302
$US_2 < \text{DLC} - \text{Hohe IK}$	76	2	78
	786	174	960

Das individuelle DLC erklärt mit knapp 72% weniger Fälle des Stoppens wie TTB. Ein möglicher Erklärungsgrund könnte sein, dass das individuelle Anspruchsniveau an Urteilssicherheit bei zwei Informationen ein anderes ist als bei einer Information. Den Personen scheinen zwei Informationen meistens zu genügen, unabhängig von der Validität der beiden Informationsquellen. Nach zwei Informationen ist es somit nicht mehr die Gesamtvalidität, die die Weitersuche bestimmt, sondern wie angenommen die Informationskonsistenz.

In Abbildung 3 wird ersichtlich, dass das Weitersuchen bzw. Stoppen der Informationssuche nicht eindeutig auf niedrige und hohe Urteilssicherheiten trennscharf kategorisiert werden kann. Zwar wird vor allem bei den niedrigen Urteilssicherheiten und niedriger Informationskonsistenz weitergesucht wird, aber bei denselben Urteilssicherheiten (vor allem im mittleren Bereich) wird häufig auch gestoppt. Das bedeutet, dass nicht einzig die Urteilssicherheiten das Stoppen bzw. Weitersuchen bestimmt.

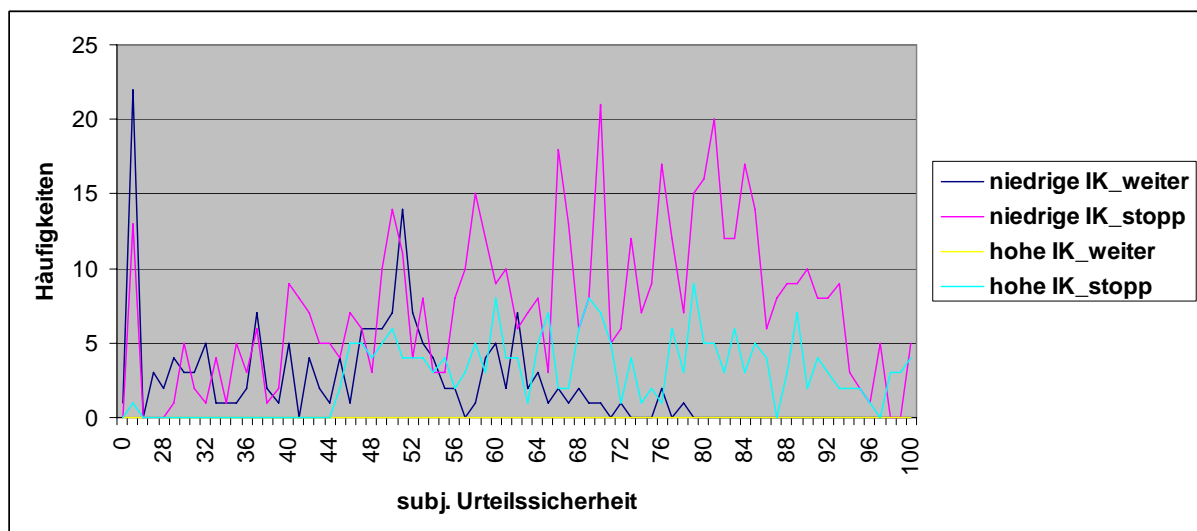


Abbildung 3: Die Häufigkeiten des Weitersuchens bzw. Stoppens der Informationssuche bei niedriger und hoher Informationskonsistenz pro subjektive Urteilssicherheit.

Diskussion

Schätzung des Desired Level of Confidence (DLC)

Es zeigt sich in diesem Experiment ein niedrigerer Mittelwert der individuellen DLCs: lag der Mittelwert bei Hausmann bei .75, finden wir in dieser Erhebung einen gemittelten Schätzwert von .55. Einerseits könnte eine mögliche Einflussgrösse für diese Mittelwertsverschiebung das leicht abgeänderte Schätzverfahren des DLCs sein. Andererseits aber könnte auch die unterschiedliche Anordnung des experimentellen Settings ein Grund dafür sein, da in diesem Experiment die Validitäten aller sechs Experten angegeben wurden und der beste Experte ausgewählt werden konnte, statt wie in den vorherigen Experimenten die Experten sequentiell vorgegeben wurden. Auffällig ist, dass ungefähr ein Viertel der Versuchspersonen in diesem Setting dann auch ein konsequentes *One Reason Decision Making* betreibt, bei Hausmann waren dies nur einzelne Personen. Hausmann & Läge (2005) fanden zwar, dass Personen ihr individuelles Anspruchsniveau je nach Gewinn anpassten, bei unterschiedlicher Anzahl von Optionen jedoch nicht. Es stellt sich somit weiterhin die Frage, wie stabil die individuellen DLCs im Grunde sind und wie abhängig das DLC von verschiedenen situativen Umwelten

(z.B. experimentelles Setting), wie auch verschiedenen Risikosituationen (z.B. Wichtigkeiten der Entscheidungskonsequenz) ist.

Bewertung der Informationen

Werden die Versuchspersonen mit einer Information konfrontiert, dann bewertet die Mehrzahl der Personen (24 von 30) die subjektive Urteilssicherheit in der Tendenz ähnlich der objektiven Validität. Dieser Befund ist jedoch nicht weiter erstaunlich, da es sich nur um eine Information handelt und wir daher nichts anderes erwartet haben. Spannender und gleichzeitig schwierig für diese Auswertungen war, dass die Validitäten sowohl unter wie auch über den individuellen DLCs gleichmässig unterschätzt werden und die Cue-Validitäten den Personen nicht die Sicherheit geben, die sie theoretisch sollten. Um dieses Problem der massiven Unterschätzung der objektiven Validitäten aufzuheben, wurden deshalb bei den weiteren Auswertungen die subjektiven Urteilssicherheiten statt der Validitäten der Personen als Sicherheitswerte der Information verwendet. Dieser Befund der Unterschätzung der Validitäten sollte eventuell bei der Schätzung der individuellen DLCs, die bisher auf den Validitäten der Informationen basiert, mitberücksichtigt werden.

Es wurde angenommen, dass bei den Validitäten unterhalb dem individuellen Anspruchsniveau die Urteilszufriedenheit tiefer liegt, als die subjektive Urteilssicherheit vermuten liesse und aber dann nach Erreichen des DLCs die Urteilszufriedenheit höher eingeschätzt wird als die Urteilssicherheiten. Wie bereits erwähnt wurde für die Testung dieser Hypothese statt der Validitäten die subjektiven Urteilssicherheiten über bzw. unterhalb des individuellen DLCs verwendet. Nimmt man die Fälle bei welchen die Urteilssicherheit des ersten konsultierten Cues unter dem individuellen DLC liegen, wird die Urteilszufriedenheit höher bewertet und bei den Urteilssicherheiten über dem DLC niedriger. So zeigen die Daten zwar einen Unterschied zwischen der subjektiven Urteilssicherheit und der Urteilszufriedenheit bei den Informationen über und unter dem DLC, jedoch zeigt der gefundene Unterschied genau in die entgegengesetzte Richtung wie angenommen und ist nicht sehr gross im Effekt. Warum einige Personen zufriedener sind, wenn die Information ihr Anspruchsniveau an Sicherheit noch nicht deckt und unzufriedener sind mit einer Information, die über ihrem individuellen Sicherheitsanspruch liegt, bleibt dennoch ungeklärt. Eventuell hat dieser Befund etwas mit der ungleichen Verteilung der subjektiven Urteilssicherheit und Urteilszufriedenheit zu tun. Da wir die Ungleichheiten in der Skalenbenutzung durch die z-Transformation versuchten auszugleichen, hat dieses Vorgehen entweder einen möglichen Effekt ausnivelliert oder aber verstärkt.

Hingegen bestätigt der Befund, dass die Personen, welche die Suche nach einer Information stoppen, zufriedener mit der Informationslage sind als diejenigen, die weitersuchen, unsere letzte Hypothese. Somit finden wir eine unterschiedliche Bewertung der Urteilszufriedenheit bei einer Information, je nach dem ob weitergesucht oder gestoppt wurde.

Bei der Bewertung von zwei Informationsquellen zeigt sich, dass die Korrelationen zwischen der A-posteriori-Validität und der Urteilssicherheit bzw. Urteilszufriedenheit verhältnismässig gering sind und die Validitäten ebenfalls beim Sicherheitsrating massiv unterschätzt werden. Da bei der A-posteriori Berechnung von zwei sich bestätigenden Informationen die subjektive

Urteilssicherheit weniger mit dieser berechneten Validität korreliert als dies bei den widersprüchlichen Informationspaaren der Fall ist, könnte vermutet werden, dass die Berechnung der A-posteriori-Validität nicht der Wahrnehmung der Gesamtvalidität der beiden Experten entspricht. Diese Berechnung müsste daher sicherlich nochmals kritisch beleuchtet werden.

Dafür zeigt sich in Phase 2 einen eindeutigen Einfluss der Informationskonsistenz auf die Urteilszufriedenheit einerseits und von der Urteilszufriedenheit auf das Stoppverhalten andererseits. Die Personen sind bei bestätigender Information zufriedener und stoppen eher und bei widersprüchlicher Information unzufriedener und suchen eher weiter. Diese beiden Resultate bestätigen unsere anfänglichen Hypothesen, dass die Informationskonsistenz bei zwei Informationen ein zentraler Faktor der Urteilszufriedenheit und die Urteilszufriedenheit ein zentraler Faktor des Stoppverhaltens ist.

Informationssuche und Stoppen der Informationssuche

Bei der Suche von einer Information verfolgen die Personen eine lexikografische Suchstrategie, d.h. sie beginnen in der Regel mit der Befragung des validesten Experten. Stehen nun aber immer zwei Experten zur Verfügung, dann löst sich dieses Muster etwas auf und es wird nicht so häufig der valideste Experte konsultiert. Diese Befund kann auf zwei Arten interpretiert werden. Entweder einige Personen betreiben nicht nur eine Entscheidung gemäss TTB, indem sie sich immer auf den validesten Experten berufen, sondern ziehen in der Konsultation des Expertenpaares ab und zu auch die Validität des zweiten Experten bei, so dass auch mal ein Expertenpaar mit zwei eher mittelmässigen Validitäten statt einem Paar mit einer hohen und einer niedrigen Validität als erstes konsultiert wird. Oder aber die Gesamtvalidität der beiden Experten wird – analog zur Interpretation der unterschätzten A-posteriori-Validität – anders verrechnet als wir das bei der A-priori-Validität angenommen haben. Anhand unserer Berechnung sollte die A-priori-Validität eines Expertenpaares immer der Validität des validesten Experten entsprechen. Es kann anhand dieser Daten nicht eindeutig darauf geschlossen werden, ob die Versuchspersonen die Wahl des Expertenpaares bloss auf die A-priori-Validität stützen, also auf eine Information à la Take The Best oder die beiden Informationen integrieren oder womöglich sogar eine A-posteriori-Validität antizipieren. Es könnte durchwegs sein, dass auch die A-posteriori-Validität, unter der Annahme von widerspruchsfreien Informationen eine Rolle bei der Informationssuche von einem Expertenpaar spielt. Insbesondere weil in diesem Experiment die Daten so konstruiert waren, dass die Expertenpaare mit einer sehr hohen Validitäten meist auch widersprüchliche Informationen zeigten, könnte dies eine Erklärung für eine im Experiment gelernte Präferenz von weniger validen Expertenpaaren sein.

Das Stoppen bzw. Weitersuchen nach einer ersten Information kann am besten durch das Anspruchsniveau an Urteilssicherheit erklärt werden. Diese bessere Erklärung vom Stoppverhalten anhand des DLC-Modells anstatt des Stopppregel von TTB ist jedoch nicht weiter erstaunlich, da das DLC aus denselben Daten ermittelt wurde, welche auch erklärt werden. Da jedoch schon Hausmann und Läge (2005) zeigen konnten, dass das DLC-Modell eine Mehrzahl der Fälle von *One Reason Decision Making* und *More Reason Decision Making* korrekt vorhersagen kann, war dies nicht Ziel dieser Untersuchung. Anders sieht es aber beim Stoppen nach zwei Informationen aus. Es scheint, dass die meisten Versuchspersonen nach zwei

Informationen genügend Sicherheit haben – unabhängig von der Validität der beiden Experten – und meist nach den beiden Informationen stoppen. Hier kann nun die tiefe Informationskonsistenz von zwei Informationen und weniger das nicht erreichte Anspruchsniveau an Urteilsicherheit die Fälle von Weitersuche erklären.

Es müsste in diesem Zusammenhang jedoch noch die Frage aufgeworfen werden, ob in Phase 2 so häufig gestoppt wird, weil zwei Informationen den Sicherheitsbedarf decken oder weil weitere konfligierende Informationen antizipiert und daher vermieden werden möchten.

Schlussbemerkung

Schon bei einer einzigen Information scheint die Urteilszufriedenheit einen zusätzlichen Erklärungswert mitzubringen, so dass die Urteilssicherheit vermutlich nicht genügt, um das Such- und Stoppverhalten bei einer Information zu erklären. Nun zeigte sich aber, sowohl in dieser Auswertung wie schon in früheren, dass einige Personen ungern auf weitere Informationen verzichten und daher häufig nach einer ersten Information noch weitere Informationen suchen. Dieses Weitersuchen nach einer Information kann das DLC sehr gut erklären, nicht aber nach zwei Informationen.

Das Verhalten beim Vorliegen von zwei Informationen weist darauf hin, wie Personen mit Informationen umgehen, falls sie nach der ersten Information weitersuchen würden. Eindeutlich zeigt sich bei zwei Informationen, dass die Informationskonsistenz eine Rolle bei der Urteilszufriedenheit spielt. Die Personen sind zufriedener, wenn die beiden Informationen einander bestätigen, was sie wiederum die Informationssuche stoppen lässt. Das bedeutet, dass nach zwei Informationen die Sicherheit gedeckt ist, da zwei Informationen unabhängig von ihrer Validität dieses Bedürfnis zu decken scheinen. Vielmehr ist es die Zufriedenheit mit der Informationslage, die nach einer Auflösung des Konfliktes von zwei unterschiedlichen Tipps verlangt. Dies ist insofern wenig erstaunlich, kennen wir doch vom Alltag die mühsamen Dilemmata in Bezug auf Entscheidungen.

Unsere Befunde zeigen, dass das Konzept der Urteilszufriedenheit einen Aspekt in der prädeziSIONalen Phase der Entscheidung abdeckt, welcher von der Validität der Information und der subjektiven Urteilssicherheit nicht erklärt werden kann. Genauer zu untersuchen bleibt unter anderem die Interaktion der Validitäten von mehr als zwei Cues mit einer kontinuierlichen Informationskonstanz und deren Einfluss auf die Urteilszufriedenheit, um den Prozess vor der Entscheidung und damit auch die Entscheidung selbst besser verstehen zu können.

Literatur

- Bröder, A. (2000a). Assessing the empirical validity of the "Take The Best"-heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1332-1346.
- Bröder, A. (2000b). "Take the best - Ignore the rest" Wann entscheiden Menschen begrenzt rational? Lengerich: Pabst.
- Bröder, A. (2005). Entscheiden mit der "adaptiven Werkzeugkiste". Ein empirisches Forschungsprogramm. Lengerich: Pabst Science Publishers.
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 115-131.
- Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional Differences in Cognitive Motivation: The Life and Times of Individuals Varying in Need for Cognition. *Psychological Bulletin*, 119(2), 197-253.
- Czerlinski, J., Gigerenzer, G., & Goldstein, D. (1999). How good are simple Heuristics? In G. Gigerenzer, P. Todd & A. R. Group (Eds.), *Simple Heuristics that make us smart*. New York, Oxford: Oxford University Press.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale *Journal of Personality Assessment*, 49, 71-75.
- Gigerenzer, G., & Goldstein, D. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650-669.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 506-528.
- Gigerenzer, G., Todd, P., & Group, A. R. (Eds.). (1999). *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.
- Hausmann-Thürig, D. (2004). Informationssuche im Entscheidungsprozess: Die Nützlichkeit von Hinweis-Cues und der Anspruch an Urteilssicherheit. Unpublished Promotionsschrift, Universität Zürich, Zürich.
- Hausmann, D., Christen, S., & Läge, D. (2005). Und sie lernen es doch! Das Einschätzen der Validität von probabilistischen Hinweis-Cues und der Einfluss der subjektiven Repräsentation auf das Urteilsverhalten. *AKZ-Forschungsbericht 8*, Zürich: Angewandte Kognitionspsychologie.
- Hausmann, D., & Läge, D. (2005a). Der Wunsch nach Sicherheit: Wie das Anspruchsniveau an Urteilssicherheit in Entscheidungssituationen unter Unsicherheit das Informationssuchverhalten steuert. *AKZ-Forschungsbericht 9*, Zürich: Angewandte Kognitionspsychologie.
- Hausmann, D., & Läge, D. (2005b). Wenn Limiten zum Strategiewechsel führen: Wechsel der Suchstrategie unter situativen Einflüssen. *AKZ-Forschungsbericht 6*, Zürich: Angewandte Kognitionspsychologie.
- Hogarth, R. M., & Karelaia, N. (2006). "Take-the-Best" and other simple strategies: Why and when they work "well" with binary cues. *Theory and Decision*, 61, 205-249.
- Newell, B., Rakow, T., Weston, N., & Shanks, D. (2004). Search strategies in decision making: The success of 'success'. *Journal of Behavioral Decision Making*, 17, 117-137.
- Newell, B., Weston, N., & Shanks, D. (2003). Empirical tests of a fast and frugal heuristic : not everyone "takes-the-best". *Organizational Behavior and Human Decision Processes*, 91, 82-96.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press.
- Roetheli, C. (2002). Wann wissen wir genug? Stopp-Mechanismen bei komplexen Entscheidungen. Universität Zürich, Zürich.
- Verplanken, B. (1993). Need for Cognition and external information search: responses to time pressure during decision-making. *Journal of Research in Personality*, 27(3), 238-252.
- Zurbriggen, S., Christen, S., Hausmann, D., & Läge, D. (2008a). Adaptivität der Informationssuche in statischen und in dynamischen Umwelten. *AKZ-Forschungsbericht 46*, Zürich: Angewandte Kognitionspsychologie.
- Zurbriggen, S., Christen, S., Hausmann, D., & Läge, D. (2008b). Entscheiden in zeitlich dynamischen Umwelten: Suchen und Stoppen mit unterschiedlich stabilen Cues. *AKZ-Forschungsbericht 44*, Zürich: Angewandte Kognitionspsychologie.
- Zurbriggen, S., & Läge, D. (2007). Ressource Zeit statt Geld: Ein Vergleich zweier Szenarien zur Entscheidungsforschung. *AKZ-Forschungsbericht 45*, Zürich: Angewandte Kognitionspsychologie.

Anhang: Exkurs A-priori-Validität versus A-posteriori-Validität

A-priori-Validität zweier Informationsquellen

Im Folgenden soll die Berechnung von der A-priori-Validität (V_{apriori}) zweier binärer, unkorrelierter Informationsquellen dargestellt werden. Beide Informationsquellen weisen jeweils auf eine von vier möglichen Optionen hin. Wir sprechen von einer hohen Informationskonsistenz, wenn beide Informationsquellen auf dieselbe Option hinweisen, d.h. wenn die beiden Hinweise sich gegenseitig bestätigen (b), und von einer niedrigen Informationskonsistenz, wenn sie sich widersprechen (w). Die beiden Informationsquellen entsprechen im Experiment zweier Experten, die Tipps mit einer bestimmten Vorhersagerate (=Validität) abgeben. Im Folgenden ist die Validität des ersten Experten = V_1 und die des zweiten Experten = V_2 , wobei die Tipps, welche sie abgeben der Versuchsperson noch nicht vorliegen.

Die A-priori-Validität (V_{apriori}) eines Expertenpaares kann aus der Summe zweier Produkte gebildet werden. Das erste Produkt ist gleich der Wahrscheinlichkeit, dass die beiden Experten identische Tipps liefern, d.h. sich in ihren Aussagen gegenseitig bestätigen (p_b), multipliziert mit der Gesamtvalidität die die beiden Experten aufweisen, sofern sie identische Hinweise liefern (V_b). Das zweite Produkt ist gleich der Wahrscheinlichkeit, dass sich die beiden Experten gegenseitig widersprechen (p_w), multipliziert mit der Validität des ersten Experten unter der Bedingung sich widersprechender Hinweise (V_{wr1}). V_{wr1} entspricht also der Wahrscheinlichkeit, dass der erste Experte recht hat (r_1), unter der Bedingung, dass beide Experten unterschiedliche Aussagen machen (w).

$$V_{\text{apriori}} = p_b * V_b + p_w * V_{\text{wr1}}$$

Hohe Informationskonsistenz

Die Validität V_b entspricht der Wahrscheinlichkeit zwei richtige Tipps von den beiden Experten zu erhalten (p_r), dividiert durch die Auftretenswahrscheinlichkeit zweier identischer Tipps (p_b), wobei $p_r = V_1 * V_2$.

$$V_b = p_r / p_b$$

Niedrige Informationskonsistenz

Die Validität V_{wr1} lässt sich berechnen aus der Wahrscheinlichkeit, dass der erste Experte einen richtigen und der zweite einen falschen Tipp abgibt (p_{wr1}), dividiert durch die Wahrscheinlichkeit dass die beiden Experten zwei sich widersprechende Tipps liefern (p_w). Es gilt anzumerken, dass es sich bei V_{wr1} nicht um einen symmetrisch berechenbaren Wert handelt, da sich die Hinweise der beiden Informationsquellen widersprechen und nur maximal einer von beiden korrekt sein kann. V_{wr1} ist dementsprechend eine bedingte Validität, nämlich die Gesamtvalidität der beiden Experten unter der Bedingung, dass die Aussage des ersten Experten korrekt ist (r_1).

$$V_{\text{wr1}} = p_{\text{wr1}} / p_w$$

$$p_w = p_{\text{wff}} + p_{\text{wr1}} + p_{\text{wr2}}$$

Dabei ist p_w gleich der Summe der Eintretenswahrscheinlichkeiten der folgenden drei Fälle:
1) Beide Experten liefern unterschiedliche, aber falsche Tipps (p_{wff}); 2) Experte 1 tippt auf die

richtige Option, Experte 2 hingegen auf die falsche (p_{wr1}); 3) Experte 1 tippt auf die falsche, Experte 2 hingegen auf die richtige Option (p_{wr2}).

$$p_{wff} = 2/3 * (1-V1) * (1-V2)$$

$$p_{wr1} = V1 * (1-V2)$$

$$p_{wr2} = V2 * (1-V1)$$

Ersetzt man in der Formel für die A-priori-Validität V_b und V_{wr1} durch die obenstehenden Formeln, so ergibt sich als A-priori-Validität V_{aprior} die Validität des ersten Experten:

$$\begin{aligned} V_{aprior} &= p_b * V_b + p_w * V_{wr1} \\ &= p_b * p_r / p_b + (p_{wff} + p_{wr1} + p_{wr2}) * p_{wr1} / (p_{wff} + p_{wr1} + p_{wr2}) \\ &= p_r + p_{wr1} = V1 * V2 + V1 * (1-V2) = V1 \end{aligned}$$

Zwei Beispiele, von zwei Tipps mit hoher Informationskonsistenz und zwei Tipps mit niedriger Informationskonsistenz sollen die Berechnung der A-priori-Validität illustrieren.

Bei einem ersten Beispiel für hohe Informationskonsistenz von zwei Tipps, wobei $V1 = .70$ und $V2 = .70$ ist, beträgt die A-priori-Validität:

$$\begin{aligned} V_{aprior} &= .49 + .21 = .70 \\ p_r &= .49; p_b = .52; p_w = .48; p_{wr1} = .21; p_{wr2} = .21; p_{wff} = .03 \\ V_b &= .49 / .52 = .94; V_{wr1} = .21 / .48 = .44 \end{aligned}$$

Bei einem zweiten Beispiel für niedrige Informationskonsistenz von zwei Tipps, wobei $V1 = .80$ und $V2 = .40$ ist, beträgt die A-priori-Validität:

$$\begin{aligned} V_{aprior} &= .32 + .48 = .80 \\ p_r &= .32; p_b = .36; p_w = .64; p_{wr1} = .48; p_{wr2} = .08; p_{wff} = .08 \\ V_b &= .32 / .36 = .89; V_{wr1} = .48 / .64 = .75 \end{aligned}$$

Diese Beispiele illustrieren noch einmal den Befund, dass ein zweiter Cue die Informationslage für eine Entscheidung a priori nicht verbessern kann, sondern nur a posteriori durch die unterschiedliche Informationskonsistenz einen Einfluss haben kann.

A-posteriori-Validität zweier Informationsquellen

Im Experiment können wir die A-posteriori-Validität im Voraus berechnen, weil wir festlegen, welche Information diese beiden Experten liefern werden, d.h. ob sie eine hohe bzw. niedrige Informationskonsistenz und damit einhergehend bestätigend oder widersprüchlich sein werden.

Hohe Informationskonsistenz

Sofern der zweite Experte die Meinung des ersten Experten bestätigt, erhöht dies die A-posteriori-Validität der Aussage des ersten Experten. Die A-posteriori-Validität (V_{apost}) der Aussage des ersten Experten berechnet sich wie folgt:

$$V_{apost-b} = V1V2 / [V1V2 + 1/3 * (1-V1) * (1-V2)]$$

wobei V_1 und V_2 wiederum die Validitäten der beiden Experten sind. Für den zweiten Experten ergibt sich dieselbe A-posteriori-Validität.

Niedrige Informationskonsistenz

Bei zwei widersprüchlichen Tipps der beiden Experten berechnet sich die A-posteriori-Validität (V_{apost1}) der Aussage des ersten Experten wie folgt:

$$V_{\text{apost-w1}} = V_1 * b / (V_1 * b + V_2 * a + 6/9 * ab)$$

wobei V_1 und V_2 die Validitäten der beiden Experten und a und b jeweils ein Drittel der Gegenwahrscheinlichkeiten der Expertenvaliditäten sind. Mit anderen Worten: V_1 und V_2 sind die Wahrscheinlichkeiten, mit denen die beiden Experten auf das Siegerpferd tippen und a und b sind die Wahrscheinlichkeiten, mit denen sie auf jedes der verbleibenden drei Pferde tippen. Dabei ist

$$a = (1 - V_1) / 3 \text{ und } b = (1 - V_2) / 3.$$

Die A-posteriori-Validität des zweiten Experten berechnet sich dementsprechend wie folgt:

$$V_{\text{apost-w2}} = V_2 * a / (V_1 * b + V_2 * a + 6/9 * ab).$$

Die A-posteriori-Validität ist also gleich der Wahrscheinlichkeit, dass der Tipp eines Experten trotzdem richtig ist, auch wenn ein anderer Experte einen widersprüchlichen Tipp abgibt.

Anhang: Resultate

Tabelle 3: Korrelationen zwischen der Validität (UV) und der subjektiven Urteilssicherheit (AV) und Urteilszufriedenheit (AV) nach einer Information in Phase 1 und das geschätzte DLC

VPNr	Validität und subj. Urteilssicherheit	Subj. Urteilssicherheit und Urteilszufriedenheit	Validität und Urteilszufriedenheit	DLC
1	0.92	0.87	0.92	0.58
2	0.30	0.38	0.59	0.30
3	1.00	0.98	0.98	0.34
4	0.99	0.95	0.96	0.46
5	0.92	0.86	0.83	0.70
6	0.79	0.85	0.72	0.74
7	0.96	0.87	0.84	0.28
8	0.99	0.78	0.75	0.74
9	0.99	0.89	0.88	0.70
10	0.97	0.96	0.97	0.62
11	0.92	0.83	0.76	0.76
12	0.98	0.97	0.97	0.25
13	0.96	0.81	0.84	0.86
14	1.00	0.93	0.93	0.80
15	0.87	0.86	0.85	0.54
16	0.83	0.87	0.93	0.64
17	0.69	0.52	0.42	0.25
18	0.98	0.91	0.87	0.76
19	0.26	0.85	0.14	0.36
20	0.99	0.88	0.83	0.40
21	0.62	0.71	0.62	0.72
22	0.35	-0.07	0.26	0.96
23	0.79	0.56	0.24	0.28
24	0.61	0.34	0.58	0.64
25	0.88	0.84	0.86	0.46
26	1.00	0.86	0.85	0.38
27	0.87	0.80	0.88	0.40
28	0.93	0.78	0.80	0.40
29	0.35	0.43	0.06	0.88
30	0.73	0.56	0.84	0.36

Tabelle 4: Korrelationen zwischen der A-posteriori-Validität (UV) und der subjektiven Urteilssicherheit (AV) und Urteilszufriedenheit nach zwei Information in Phase 2

VPNr	A-posteriori-Validität und subjektive Urteilssicherheit	subjektive Urteilssicherheit und Urteilszufriedenheit	A-posteriori-Validität und Urteilszufriedenheit
1	0.89	0.93	0.79
2	0.14	0.78	0.16
3	0.97	0.95	0.91
4	0.77	0.98	0.77
5	0.76	0.78	0.60
6	0.79	0.89	0.75
7	0.73	0.68	0.71
8	0.73	0.93	0.59
9	0.87	0.76	0.74
10	0.89	0.88	0.84
11	0.90	0.88	0.81
12	0.90	0.95	0.92
13	0.92	0.73	0.68
14	0.85	0.84	0.71
15	0.81	0.77	0.80
16	0.78	0.87	0.68
17	0.22	0.56	0.17
18	0.96	0.88	0.85
19	0.18	0.13	0.49
20	0.76	0.93	0.73
21	0.19	0.42	0.51
22	0.29	0.04	-0.10
23	0.79	0.56	0.38
24	0.25	0.40	0.57
25	0.80	0.86	0.72
26	0.86	0.95	0.86
27	0.57	0.72	0.44
28	0.88	0.93	0.83
29	0.47	0.45	0.27
30	0.65	0.85	0.56

LERNUMWELTEN UND CUE-AUSPRÄGUNGSRATE

Kapitel 2: Wie lernen wir zu diskriminieren?

Einleitung

Welche Stadt ist grösser: Chongqing oder Shenyang? Chongqing ist die grössere Stadt, denn sie beherbergt zurzeit ungefähr 7.7 Millionen Einwohner und ist somit die viertgrösste Stadt Chinas. Shenyang ist die fünftgrösste Stadt Chinas mit circa 7 Millionen Einwohnern. Gäbe es kein Wikipedia oder Statistiken mit diesen Einwohnerzahlen, müsste man anhand von verschiedenen Hinweisen ausfindig machen, welche Stadt wohl die grössere ist, gegeben dem Fall, dass man keine der beiden Städte erkennt. Gigerenzer und die ABC- Research Group (1991) argumentieren, basierend auf der Theorie vom probabilistischen mentalen Modell (PMM), dass Menschen in der Regel nicht die Städtegrössen, also die Zielvariable, lernen, sondern die Hinweise, welche etwas über diese Zielvariable aussagen. Das PMM wurde von Gigerenzer et al. basierend auf dem Linsenmodell von Brunswik (1952) entworfen. Die Hinweise werden als Cues definiert und ihre Güte als Validität. "If subjects had repeated experience with a reference class, a target variable, and cues in their environment, we assume that cue validities correspond to the ecological validities" (1991, S.510). Im PMM werden die Cue-Validitäten im Bezug auf probabilistische Informationen im Gedächtnis definiert: „A PMM connects the specific structure of the task with a probability structure of a corresponding natural environment (stored in long-term memory)" (S.507, 1991) und "In Brunswik's view, cue validities are learned by observing the frequencies of co-occurrences in an environment" (1991, S.510.). Das bedeutet, dass im Gedächtnis gespeichert wird, welche Hinweise mit welchen Zielvariablen korrelieren, wie gut diese Cues dies tun und wie häufig dies vorkommt. Diesbezüglich ist die Informationssuche theoretisch als ein Informationsabruf im Gedächtnis gemäss den gelernten Cue-Validitäten konzipiert

Gigerenzer et al. haben in den vergangenen Jahren eine Reihe von Heuristiken in der „Adaptiven Toolbox“ (Gigerenzer et al. 1999)¹ formuliert, die Simons Argumentation einer begrenzten Rationalität (Simon, 1955)² Rechnung tragen und dabei zumindest potentiell so gebaut sind, dass sie eine Phase der Informationssuche einschliessen (Gigerenzer & Goldstein, 1996; Gigerenzer et al., 1991). Es wird davon ausgegangen, dass die Informationen nur solange gesucht und verarbeitet werden, bis eine ausreichend gute Basis für die Entscheidung vorliegt. Diese Suchstrategie nach einer guten Information nennt man auch *One Reason Decision Making* (ORDM). Die berühmteste Heuristik aus der Adaptiven Toolbox ist die *Take the Best* Heuristik (TTB). "«Take The Best» first tries the cue with the highest validity, and if it does not discriminate, the next best cue, and so on. Its motto is «take the best, ignore the rest»"

¹ Zur Diskussion von subjektiver Rangreihe von Cues und deren Zusammenhang zur ökologischen Validitätenrangreihe siehe Schriftenwechsel von Dougherty et al. und Gigerenzer et al. in der Zeitschrift *Psychological Review* (Dougherty, Franco-Watkins, & Thomas, 2008; Dougherty, Rick, & Franco-Watkins, 2008; Gigerenzer, Hoffrage, & Goldstein, 2008).

² In dem Artikel von 1955 wird der Begriff *bounded rationality* noch nicht verwendet, aber die gedanklichen Bausteine der neuen Anti-ökonomischen Theorie nahmen dort in der Argumentation, dass der Mensch nicht mit einem Computer zu vergleichen sei, sondern sein Wissen und seine Fähigkeiten limitiert seien, ihren Anfang.

(Gigerenzer, Todd, & ABC-Research Group, 1999, S.81). Neben der Suchregel wird eine Heuristik in der Adaptiven Toolbox auch durch die Stopp- und Entscheidungsregel definiert.

Bisherige experimentelle Verhaltensbefunde zeigen, dass Personen ihre Informationssuche nach der Validität ausrichten: z.B. fanden Rieskamp und Hoffrage (1999), dass die Versuchspersonen unter Zeitdruck die valideren Cues länger anschauten als die weniger validen Cues und demzufolge die Zeit, bei einem Cue zu bleiben, mit dessen Validität hoch korreliert. In einem späteren Experiment (Rieskamp & Hoffrage 2008) fanden die Autoren, dass Personen ihre Suchstrategien adaptiv an die Situation anpassten, d.h. bei hohem Zeitdruck Heuristiken wie die lexikografische *Take The Best* verwendeten. Die Autoren liessen die Validitäten in diesen Experimenten jedoch nicht lernen, sondern erklärten das Konzept den Versuchspersonen und gaben die Validitäten der Cues an. Da dies jedoch nicht der ursprünglichen Theorie von Informationssuche im Gedächtnis entspricht, liess Bröder (2000) in einem Experiment die Validitäten von fünf Cues anhand einer ausgedehnten Lernphase (250 Durchgänge für 5 Cues) lernen (siehe Abbildung 4), wobei die Cues immer im Paarvergleich gelernt wurden. Da Bröder jedoch die Entscheidungsregeln untersuchte und nicht die Suchregeln, gab er die Rangreihen der Cues bei der Testphase vor. Die Befunde zu gelernten Validitäten zeigten, dass die Personen die objektiven Rangreihe der Cues gut einschätzen konnten. Bei der Entscheidungsregel fand Bröder grosse individuelle Differenzen und schlägt deshalb einen idiographischen Ansatz für die Entscheidungsforschung vor.

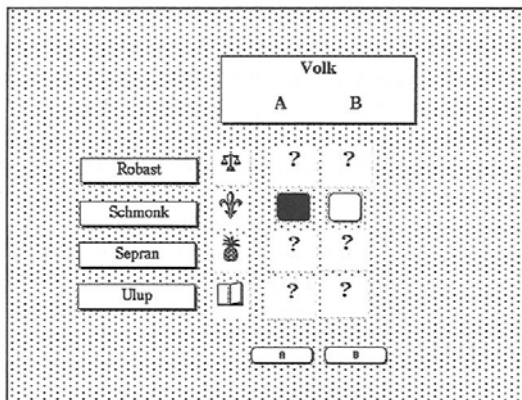


Abbildung 4-1: Ein Durchgang der Lernphase.

Abbildung 4: Lernphase aus Bröder (2000)

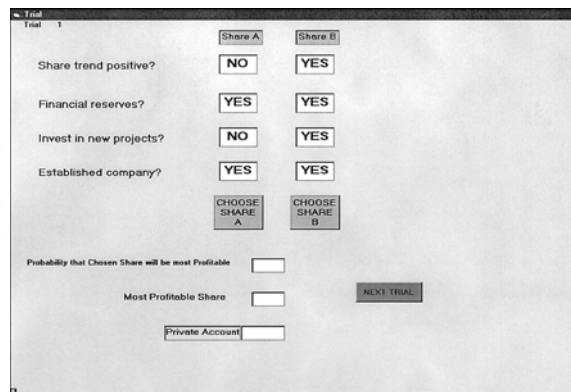


Abbildung 5: Lernphase aus Newell und Shanks (2003)

Weil das Lernen der Cue-Güte im experimentellen Setting sehr aufwändig ist, ging Bröder dazu über, den Versuchspersonen die Rangreihe der Cue-Validitäten jeweils vorzugeben, wie z.B. im Börseninvestitions-Szenario (Bröder, 2000). Dieses Börseninvestitions-Szenario wurde von Newell und Kollegen übernommen (siehe Abbildung 5), um ebenfalls die Validitäten von Cues im Paarvergleich lernen zu lassen.

Newell und Shanks (2003) fanden dabei, dass die Versuchspersonen, nachdem die Validitäten in einer ersten Phase gelernt wurden, diese Reihenfolge auch in der aktiven Informationssuche einhielten – zumindest auf Gruppenniveau. Auf individuellem Niveau zeigte sich eine grosse Verhaltensvarianz, weshalb die Autoren dann in einem zweiten Experiment die Reihenfolge der Validitäten den Versuchspersonen vor der Testphase anhand von Tipps vorgaben.

Ein Cue ist aber vor allem dann nützlich, wenn er eine hohe Validität besitzt und – in einem Zwei-Alternativen-Fall – zwischen den beiden Optionen diskriminiert. Die Diskriminationsrate d_i des i -ten Cues definiert sich gemäss Gigerenzer (1999, S.85) aus der Häufigkeit, mit welcher dieser Cue zwischen einem Paar von zwei Objekten einer Referenzklasse (N) diskriminiert (wobei x und y die binären Ausprägungen des Cues sind):

$$d_i = 2x_i y_i / 1 - (1/N)$$

Die Kombination der beiden Gütemasse Validität und Diskriminationsrate wurde als Mass der Nützlichkeit der Cues von Martignon und Hoffrage (1999) als Success formuliert. Success (S) beschreibt einerseits die Suchregel, die lexikografisch gemäss einer Multiplikation von Validität und Diskriminationsrate vorgeht, andererseits aber auch die Stoppregel: nach einem konsultierten Cue muss die Informationssuche gestoppt werden, falls der Cue nicht diskriminiert, muss geraten werden. Folgende Formel gilt somit für Success (d = Diskriminationsrate und v = Validität des Cues):

$$S = d*v + ((1 - d)*0.5)$$

Newell, Rakow, Weston und Shanks (2004) weiteten das Lernen auf die Diskriminationsrate aus, wobei die Reihenfolge der Cues nicht mehr vorgegeben wurde, sondern sie die Versuchspersonen in 64 Runden die Validitäts- und Diskriminationsrate im bekannten Börsenszenario lernen liessen. Die Fragebogendaten deuteten dabei ebenfalls darauf hin, dass die Versuchspersonen fähig sind, die Rangreihe der Cues gemäss Usefulness³, Validität und Diskriminationsrate deklarativ anzugeben, und gemäss Success auf den „richtigen“ Cue zu tippen. Bei den numerischen Einschätzungen wurde die Validitätsrate jedoch systematisch unterschätzt (von .09 bis .27), die Diskriminationsrate überschätzt (.04 bis .16). Das Fazit dieser Forschungen war dann auch: Können die Personen die Reihenfolge der Cues frei wählen, dann setzt sich die Suchregel Success als bestes Erklärungsmodell durch. Dabei fand sich sowohl bei den Suchstrategien als auch beim Stoppverhalten grosse individuelle Differenzen. Obwohl das Stoppverhalten nicht im Fokus dieser Untersuchung stand, zeigte sich, dass nur gerade 1% der Personen gemäss Success stoppte. Somit stellt sich die Frage, wie denn die Versuchspersonen weitersuchten, wenn sie die Suche nach dem ersten Cue nicht beendeten.

Christen, Hausmann & Läge (2006) entwickelten das sogenannte Äffchen-Experiment (siehe Abbildung 6), in welchem die Validität und die Diskriminationsrate nicht anhand paarweiser Cues, sondern anhand von Einzelinformationen gelernt werden konnten (Kriteriumsabhängig in 100 bis 200 Durchgängen).

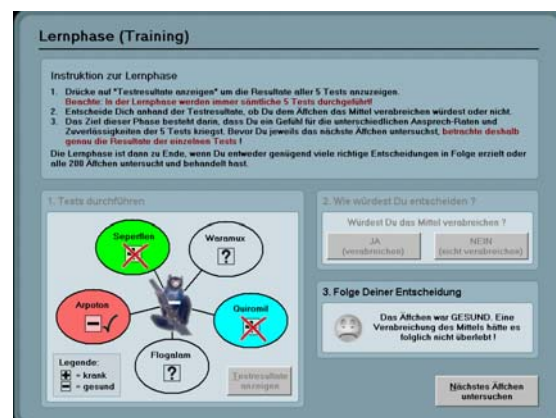


Abbildung 6: Lernphase aus Christen et al. (2006)

³ Usefulness wurde bei den Autoren im Sinne von der allgemeinen Nützlichkeit eines Cues verwendet und mit Success gleichgesetzt. Hoffrage, Hertwig, & Czienskowski (2003) definierten Usefulness jedoch als die Kombination von $V*D$, jedoch ohne die Stoppregel eines One-Reason Decision Making. Läge, Hausmann, Christen, & Daub (2005) überprüften diese beiden Masse experimentell.

Das Setting des Experimentes unterschied sich somit fundamental von den bisherigen Experimenten: es handelte sich nicht um einen Zwei-Alternativen-Fall. Das heisst, es wurde keine Wahl zwischen zwei Alternativen erzwungen, sondern es musste jeweils entschieden werden, ob einem Äffchen ein Medikament verabreicht werden sollte oder nicht. Verschiedene Tests konnten durchgeführt werden, wobei ein Test, d.h. ein Cue mit einer gewissen Trefferrate anzeigte, ob das Äffchen krank war oder nicht. Dementsprechend war auch die Diskriminationsrate dadurch definiert, ob ein Cue einen Hinweis lieferte oder nicht und nicht wie bis anhin, ob ein Cue zwei gleiche bzw. zwei verschiedene Informationen zu den zwei zur Verfügung stehenden Alternativen lieferte. In der Testphase mussten die Personen die Tests ganz spezifisch nach ihrer Validität, Diskriminationsrate oder Success-Rate einsetzen. Die Befunde zeigten, dass die Mehrzahl der Personen fähig sind, die Gütemasse durch Beobachtungslernen erwerben zu können, und diese Masse auch korrekt einsetzen.

Im Grossen und Ganzen zeigen die bisherigen Befunde, dass die experimentelle Lernbarkeit der Cue-Güte Validität und Diskriminationsrate machbar, aber nicht sehr einfach ist und dass grosse individuelle Unterschiede in der Anwendung der Such-, Stopp- und Entscheidungsregel bestehen. Bei allen bisherigen berichteten Experimenten wurde nicht untersucht oder nicht berichtet, wie die Informationslage aussah, wenn nach einer Information weitergesucht wurde oder eben nicht. Wir wissen zwar, dass die Umwelt, z.B. die Informationskosten das Stoppen der Suche beeinflussen (z.B. Läge, Hausmann, & Christen, 2005; Rakow, Newell, Fayers, & Hersby, 2005), aber auch verschiedenen individuelle Faktoren, wie z.B. das Anspruchsniveau an Urteilssicherheit (Hausmann & Läge, 2008) oder die Intelligenz (Bröder, 2003) bestimmt, ob eine Person weiter Informationen sucht oder sich mit einer Information zufrieden gibt. Dabei ist die Frage ungeklärt, wie die Umwelt und die schon vorhandenen Informationen die Informationssuche beeinflusst, wenn nicht gleich nach einer Information gestoppt wird. Wir gehen davon aus, dass die Cue-Suche schon bei dem zweiten Cue sehr stark streut, weil die Heuristiken nicht streng befolgt werden, sondern sich die Suche adaptiv an die Informationslage anpasst. Wie die Informationssuche aussieht, wenn der Cue nicht diskriminiert, dieser Fragestellung wurde bis anhin kaum Beachtung geschenkt. Gemäss dem validitätsbasierten Ansatz müsste bei dem nächst validesten Cue weitergesucht werden, gemäss der Stoppregel von Success dürfte überhaupt nicht weitergesucht werden. Wir nehmen an, dass im Falle einer weiterführenden Suche nach einer nicht diskriminierenden Information diese nicht diskriminierende Information einen Informationswert hat und die Informationssuche dementsprechend beeinflusst. Dazu ein Beispiel im Sinne des Städte-Paradigmas von Gigerenzer et al.

(Gigerenzer & Goldstein, 1996): Nehmen wir an, es gäbe drei Cues mit binärer Ausprägung, anhand derer nach probabilistischer Logik auf die Grösse von zwei Städten geschlussfolgert werden kann (siehe Tabelle 5): Cue 1 deutet darauf hin, ob eine Stadt einen ICE Bahnhof hat oder nicht; Cue 2 zeigt an, ob eine Stadt eine Universität hat oder nicht und Cue 3 indiziert einen Flughafen. Nehmen wir zur Vereinfachung der Fragestellung an, dass die Validitäten aller drei Cues gleich gross sind. Die Diskriminationsrate des ersten und dritten Cues würde .32 betragen und des zweiten Cues .50. Nun kann eine Diskriminationsrate von .32 auf zwei verschiedene Arten zustande kommen: z.B. könnte der Cue 1 bei 80% aller Städte vorhanden sein und bei 20% aller Städte nicht. Bei Cue 3 hingegen könnten nur 20% aller Städte einen Flughafen haben, 80% aller Städte hätten keinen. Das bedeutet, dass ein Cue neben der Diskriminationsrate auch eine Ausprägungsrate hat. Die Ausprägungsrate (AR) ist somit definiert

durch den Prozentsatz von Städten, die über eine positive Ausprägung eines Cues verfügen (bspw. 20% der Städte haben einen Flughafen => AR = .20). Aus der Ausprägungsrate eines Cues lässt sich die unspezifischere Diskriminationsrate (DR) rechnerisch eindeutig ableiten (=> DR = .32).

Tabelle 5: Hypothetische Annahme der Validität, Diskriminationsrate und Ausprägungsrate von drei Cues.

	Validität	Diskriminationsrate (DR)	Ausprägungsrate (AR)
Cue 1: ICE Bahnstation	1.00	.32	.80
Cue 2: Universität	1.00	.50	.50
Cue 3: Flughafen	1.00	.32	.20

Diese neue Theorieannahme steht im Sinne einer begrenzten Rationalität und postuliert eine Abhängigkeit der Suchstrategie von der Diskriminationsrate und der Validität beim ersten gesuchten Cue und eine Abhängigkeit der weiterführenden Suche von der Ausprägungsrate, wenn der erste Cue nicht diskriminiert. Im obigen Beispiel wäre der Cue 2 mit der höchsten Diskriminationsrate der Cue, der gemäss der Success-Strategie der beste wäre und somit als erstes gesucht werden müsste. Würde nun Cue 2 bei beiden Städten nicht diskriminieren, müsste gemäss Success unabhängig von der Information die Suche gestoppt werden. Gemäss dem validitätsbasierten Ansatz von TTB müsste nach einer diskriminierenden Information die Suche gestoppt werden und bei einer nicht-diskriminierenden der nächstvalideste Cue (Cue 1 oder 3) konsultiert werden. Wir nehmen nun an, dass die Art der nicht-diskriminierenden Information (also ein „Ja, bei beiden Städten vorhanden“ oder „Nein, bei beiden Städten nicht vorhanden“), die Suchtendenz für den nächsten Cue bestimmt. Falls man weiss, dass beide Städte eine Universität haben, dann sucht man eher nach dem Cue, der bei grossen Städten diskriminiert und eine niedrige Ausprägungsrate hat (Cue 3). Cue 3 zeigt, dass wenige Städte einen Flughafen haben, daher diskriminiert er bei grossen Städten eher als bei kleinen Städten. Haben nun aber beide Städte keine Universität, also zeigt der Cue 2 zwei negative Ausprägungen, würde man eher auf eine kleinere Stadt tippen und daher nach dem ersten Cue einen Cue konsultieren, der eher bei kleinen Städten diskriminiert. Man würde den Cue 1 mit der hohen Ausprägungsrate als zweiten Cue ansehen. Eine solche Suchstrategie lässt sich sowohl gegen die Suchstrategie- und Stoppstrategie des Success, wie auch gegen die validitätsbasierte Suchstrategie testen.

Falls sich das Konzept der Informationssuche unter Berücksichtigung der Ausprägungsrate experimentell bestätigen lässt, müsste die bisherige Theorie der streng sequentiellen Suchregel neu überarbeitet und um eine Abhängigkeit der Suchregel von Informationsgrundlage und Ausprägungsrate als moderierende Variablen erweitert werden. Falls eine Anwendung der Ausprägungsrate in der Informationssuche gefunden wird, stellt sich in einem zweiten Schritt die Frage, wie diese Ausprägungsrate gelernt wird. Im Probabilistischen Mentalen Modell wird davon ausgegangen, dass die Validitäten der Cues durch Beobachtung gelernt werden. Ist es im Erfahren der Häufigkeit von Ausprägungen analog zu Christen et al., dass man sich ein Bild über die Diskriminierfähigkeit eines Hinweises macht oder im direkten Paarvergleich analog zu der experimentellen Tradition, wie wir sie anhand der Experimente von Bröder und Newell referierten? Hier möchten wir nun ansetzen und untersuchen, auf welche Art die Validität und die Diskriminationsrate besser gelernt werden. Unsere Annahme ist, dass die Ausprägungsrate näher am hypothetischen Konstrukt eines „inneren Zählers“ ist als die Diskriminationsrate. Kennt man die Ausprägungsrate, kann die Diskriminationsrate daraus

gefolgt und angewendet werden, dazu braucht es nicht zwingend einen Paarvergleich. Konkret ergeben sich folgende Fragestellungen und Hypothesen:

1. Wird die sequentielle Informationssuche an die Ausprägungsrate angepasst?

Wie sieht die Suche nach dem ersten Cue aus? Ändert eine Person ihre Suchregel in Abhängigkeit der Ausprägung bei nicht-diskriminierender Information? Wir nehmen ein *One Reason Decision Making* an und formulieren als Hypothese, die Auswahl des ersten Cues gemäss Success. Suchen die Versuchspersonen gemäss Success, müssten sie nach dem ersten Cue die Informationssuche stoppen. Falls die Personen jedoch weitersuchen, nehmen wir an, dass die Aussage, ob die beiden Alternativen über dieselbe Cue-Ausprägungen verfügen, als Zusatzinformation zu werten ist, was die weitere Suchrichtung festlegen sollte. Bei zwei positiven Cue-Ausprägungen (+ +) sollte häufiger als zweiter Cue ein Cue mit niedriger Ausprägungsrate gewählt, als mit hoher. Bei zwei negativen Cue-Ausprägungen (- -) sollte als zweiter Cue eher ein Cue mit hoher Ausprägungsrate gewählt, als mit niedriger.

2. Ist die Anwendung der Ausprägungsrate abhängig vom Lernformat?

Anhand zweier Lerngruppen untersuchen wir den Einfluss des Lernformats auf die Informationssuche und die Anwendung der Ausprägungsrate. Die eine Lerngruppe A lernt analog zu den Experimenten von Bröder und Newell mit paarweisen Vorgaben der Cue-Informationen. Die andere Lerngruppe B lernt nur die Ausprägungsrate von den einzelnen Wahlalternativen analog zum Äffchen-Experiment von Christen et al. In der Anwendung der Ausprägungsrate sollte die Lerngruppe B (Ausprägungsrate-Lernen) sich gemäss unseren Hypothesen in Abhängigkeit der vorgegebenen nicht diskriminierenden Information (siehe 1.) verhalten, die Lerngruppe A (Diskriminationslernen) jedoch nicht.

3. Ist das Lernen der Cue- Gütemasse Validität und Diskriminationsrate abhängig vom Lernformat?

Anhand von zwei Lerngruppen untersuchen wir die Lernbarkeit der Gütemasse Validität, Diskriminationsrate und deren Anwendung beim Success und testen das Wissen sowohl implizit anhand von Verhaltenstests, wie auch explizit anhand von Nachbefragungen. Wir nehmen an, dass die beiden Lerngruppen sich nicht unterscheiden: Die Lerngruppe A (Diskriminationslernen) sollte sowohl die Validitäten, wie auch die Diskriminationsraten gut lernen. Die Lerngruppe B sollte die Validitäten ebenfalls lernen, aber etwas weniger gut und die Diskriminationsrate über die Ausprägungsrate aber genauso gut erschliessen können.

Experiment 1: Wird die sequentielle Informationssuche an die Ausprägungsrate angepasst?

Wie gehen Personen vor, wenn sie Informationen frei suchen können? Dazu muss Wissen über die Gütemasse der Informationen erworben werden. Obwohl man in der Vergangenheit dazu übergegangen ist, die Gütemasse vorzugeben, gehen wir nochmals einen Schritt zurück und lassen die Personen diese erlernen. Da es in diesem Experiment nicht primäres Ziel ist, die Suchstrategien Validität, Diskriminationsrate oder Success zu testen, sondern vielmehr die Weitersuche nach dem ersten Cue, lassen wir die Personen die Ausprägungsrate der Cues

beobachten. Gemäss unserer Annahme sollten die Personen den zweiten Cue in der Testphase in Abhängigkeit der Ausprägungsrate aufdecken. Wenn ein Cue zuvor nicht diskriminierte, sollte bei zwei positiven Cueausprägungen (+ +) häufiger als nächster Cue ein Cue mit niedriger Ausprägungsrate gewählt werden, als mit hoher Ausprägungsrate. Bei zwei negativen Cueausprägungen (- -) sollte als nächster Cue eher ein Cue mit hoher Ausprägungsrate gewählt werden, als mit niedriger Ausprägungsrate. Diese Befunde sollten sich einerseits bei der gesamten Informationssuche, andererseits aber auch bei der Konsultation nach dem ersten Cue zeigen.

Methode

Subjects

Es nahmen 24 Versuchspersonen teil (14 Männer und 10 Frauen), darunter 12 Psychologie-Studierende der Universität Zürich. Das Durchschnittsalter betrug 25 Jahre (Range 17-53). Die Psychologie-Studierende müssen im Rahmen ihrer Ausbildung an der Universität Zürich mindestens fünf Stunden an Experimenten teilnehmen, an welchen ist ihnen dabei frei gestellt. Die Nicht-Studierenden nahmen freiwillig teil. Unter der besseren Hälfte wurde eine Gewinnsumme von zweimal CHF 50.- verlost.

Apparatus

Die Lernphase wurde mit dem Powerpoint Programm erstellt, die Testphase wurde in Microsoft Visual Basic 6.0 programmiert. Das Experiment wurde vollständig auf einem PC mit Windows XP-Betriebssystem durchgeführt. Eine Nachbefragung zur Ausprägungsrate wurde auf Papier ausgefüllt.

Stimuli, Design and Procedure

Das Experiment bestand aus einer Lernphase, einer Testphase und einer Nachbefragung. Die Aufgabe in der Lernphase bestand darin, sich ein Gefühl der Nützlichkeit der Cues in Betracht der Beurteilung von Städtegrössen anzueignen. Statt die Ausprägungen paarweise zu lernen, wurden die Ausprägungen von sechs Cues einzeln pro Stadt dargeboten, wenn diese vorhanden waren, dann waren sie fett hervorgehoben (siehe Abbildung 7). Zudem wurde auch die Referenzmenge, d.h. die Cue- Ausprägungen bei der kleinsten und grössten Stadt angegeben. Diese Lernphase umfasste 60 Städte und dauerte fünf Minuten. Der Text lautete folgendermassen:

In diesem Experiment wirst Du als Managerin bzw. Manager für eine internationale Firma tätig sein, die den lukrativen chinesischen Markt erobern will. Deine Aufgabe wird es sein, eine Filiale in der jeweils grösseren von zwei Städten zu eröffnen. Bevor Du jedoch mit Deiner neuen Arbeit beginnen kannst, sponsert dir deine neue Auftragsgeberin eine Reise durch China, damit Du Dir ein Bild über die Grösse der Städte machen kannst. Denn die chinesische Regierung gibt grundsätzlich keine Kennzahlen (wie z.B. Einwohnergrössen) bekannt. Indem Du nun in diesem ersten Teil durch die verschiedenen Städte reist, erfährst Du indirekt etwas über die generelle Grösse von Städten. Denn aus dem Vorhandensein bzw. Fehlen von Hinweisen (wie z. B. einer Fahrradfabrik, eines Krankenhauses, von Linienbussen, eines Parteibüros, einer eigenen Tageszeitung oder eines Volleyball-Teams) kann mit einer gewissen Fehlerwahrscheinlichkeit auf die relative Grösse einer Stadt geschlossen werden. Die Grösse einer Stadt ist jeweils mit einem hellblauen Kreis im Vergleich zur grössten und kleinsten Stadt dargestellt. Im ersten Teil geht es nun darum, Dir ein Gefühl für die Nützlichkeit der einzelnen Hinweise im Bezug auf die Städtegrösse zu bilden. Geniesse die Reise, denn nach jeweils 5 Sekunden bist du jeweils automatisch in einer neuen Stadt.

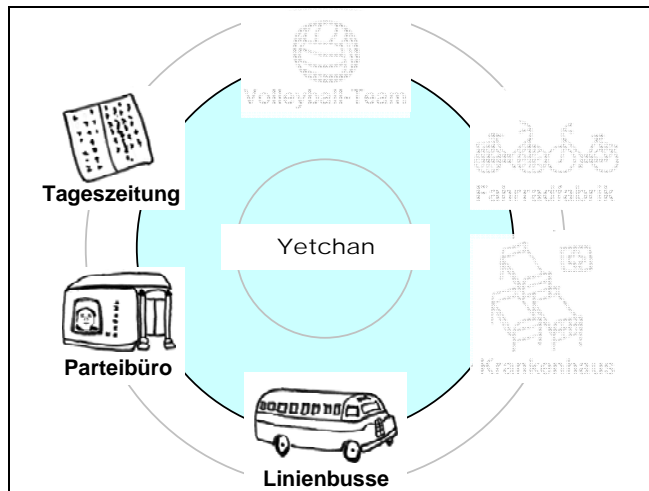


Abbildung 7: Lernkarte mit Angaben zur Grösse (türkiser Kreis) einer chinesischen Stadt und den vorhandenen Cues (fett hervorgehoben)

Die Testphase basierte auf einem Information Board (siehe Abbildung 8), analog dem Mouse-lab von Payne et al. (Payne, Bettman, & Johnson, 1988). Das Setting wurde von Hausmann und Kollegen übernommen (Hausmann, Christen, & Läge, 2005), welches dem Städteparadigma von Gigerenzer entsprach, wobei die Städte „chinesifiziert“ wurden, d.h. anhand erfundener Hinweise (dieselben sechs Cues, wie in der Lernphase) musste auf die Grösse von fiktiven, chinesischen Städte geschlossen werden und jeweils die grössere von zwei Städten ausgewählt werden. Die Testphase beinhaltete 60 Entscheidungen, wobei den Versuchspersonen 180 Informationen dafür zur Verfügung standen. Die Coverstory lautete wie folgt:

Deine Firma freut sich, dich als ihre/n Manager/in zu haben. Sie beauftragt dich 60 neue Filialen in chinesischen Städten verschiedener Regionen zu eröffnen. Die PR-Abteilung Deiner Firma gibt Dir jeweils zwei in Frage kommende Städte pro Region vor. Für die Gewinnmaximierung des Konzerns ist es von zentraler Bedeutung, dass Du die Filiale jeweils in der grösseren der beiden Städte eröffnest. Leider gibt die chinesische Regierung für ihre Städte keine Einwohnerzahlen bekannt. Auch erfährst Du von deiner Firma nicht, ob Fahrradfabrik, Krankenhaus, Linienbusse, Parteibüro, Tageszeitung oder Volleyball-Team bei diesen Städten vorhanden sind. Durch Bestechung einer chinesischen Beamtin kannst Du jedoch weiterhin an die Hinweise rankommen und dir ein Bild über die Grösse der zur Auswahl stehenden Städte machen. Hohe Bestechungskosten schmälern natürlich die Gewinnerwartung Deiner Firma. Deshalb gibt Dir Deine Firma einen fixen Geldbetrag, mit dem Du die Beamtin insgesamt 180mal bestechen kannst. Deine Aufgabe besteht darin, zu entscheiden, welche der beiden Städte die grössere ist!

Das Interface zeigt die Testphase mit der Frage: 'Welche Stadt hat mehr Einwohner?' mit den Optionen 'Zhoudinghen' und 'Tungjangshan'. Die Antwort 'Tungjangshan' ist als 'RICHTIG' markiert. Darunter befindet sich eine Tabelle zur Informationssuche:

Hinweis	Info	Zhoudinghen	Tungjangshan
Volleyball-Team	kaufen	-	+
Krankenhaus	kaufen	-	-
Tageszeitung	kaufen	+	+
Fahrradfabrik	kaufen	-	-
Linienbusse	kaufen		
Parteibüro	kaufen		

Statistiken am unteren Rand:

- Du kannst insgesamt 180 Informationen kaufen.
- Anzahl verbleibende Informationen: 174
- Pro Aufgabe musst Du mindestens 1 Information kaufen.
- Anzahl verbleibende Aufgaben: 58

Ein 'Weiter' Button befindet sich am unteren Rand.

Abbildung 8: Testphase mit sechs Cues und aktiver Informationssuche.

Bei der Nachbefragung mussten die Leute die sechs Cues im Bezug auf ihre Ausprägungsrate in Rangreihen ordnen. Zusätzlich mussten die Ikons der sechs Hinweise aus der Lernphase nach ihrer Sympathie bewertet werden, damit eine Unvoreingenommenheit gegenüber den Ikons gewährleistet werden konnte.

Die Gütemasse der Cues entsprachen den a priori festgelegten Validitäten, Diskriminationsraten und Ausprägungsraten (siehe Tabelle 6) und die Verteilung der Städtegrößen war linear.

Tabelle 6: Validität, Diskriminationsrate, Ausprägungsrate, Success und Usefulness der sechs Cues

	Validität (V)	Diskriminationsrate (DR)	Ausprägungsrate (AR)	Success	Usefulness
Cue 1: Linienbusse	.83	.20	.90	.57	.17
Cue 2: Parteibüro	.88	.40	.70	.65	.35
Cue 3: Tageszeitung	.70	.50	.50	.60	.35
Cue 4: Volleyball-Team	.85	.43	.30	.65	.37
Cue 5: Krankenhaus	.90	.17	.10	.57	.15
Cue 6: Fahrradfabrik	1.00	.10	.05	.55	.10

Gemäss diesen verschiedenen Cue-Gütemassen können unterschiedliche Suchstrategien für den ersten Cue ausgewertet werden. Wenn eine Person ihre Informationssuche nach der Validität ausrichtet, beginnt sie die Suche bei dem Cue 6 und fährt bei dem Cue 5 fort. Wäre die Diskriminationsrate das Mass der Suche, dann sollte mit dem Cue 3 begonnen werden und mit dem Cue 4 weitergesucht werden. Falls die Personen ihre Suche nach der Ausprägungsrate ausrichten, dann würden sie mit dem Cue 1 beginnen und mit dem Cue 2 weiterfahren. Wenn die Personen aber Validität und Diskriminationsrate gemäss Success kombinieren, so wie wir dies als Norm annehmen, dann sollten sie ihre Informationssuche mit dem Cue 2 oder 4 beginnen und dann Stoppen. Entspricht die Kombination von Validität und Diskriminationsrate dem Konzept der Usefulness, dann wäre Cue 4 der Favorit und die Weitersuche wäre bei Cue 2 oder Cue 3 angebracht.

Resultate

Um die sequentielle Informationssuche und ihre Abhängigkeit von nicht-diskriminierende Informationen zu analysieren, wird in einem ersten Schritt die Suchhäufigkeit bei den einzelnen Cues dargestellt. In einem zweiten Schritt wird die Abhängigkeit dieser Informationssuche im Bezug auf die zuvor aufgedeckte nicht-diskriminierende Information analysiert. Tabelle 7 zeigt, welche Cues die 24 Versuchspersonen in den insgesamt 60 Durchgängen insgesamt bevorzugten und als erstes wählten. Cue 2 war eindeutig der beliebteste Cue, sowohl über alle Durchgänge gesehen (signifikante Unterschied bei Cue 1: $t=2.2$, $p<0.5$ und Cue 6: $t=4.1$, $p<.001$), wie auch als erstes (signifikante Unterschied jedoch nur mit Cue 6, $t=2.7$, $p<.05$). In den ersten dreissig Entscheidungen suchten die Versuchspersonen weniger Informationen als in den letzten dreissig Durchgängen: in der ersten Hälfte kaufte jede Person im Durchschnitt 2.7 Cues und in der zweiten Hälfte 3 Cues. Dieser Unterschied ist jedoch nicht signifikant ($t(23) = -1.9$; $p = .06$).

Tabelle 7: Häufigkeit der gesamten Informationssuche und des ersten Cues.

	Cue 1	Cue 2	Cue 3	Cue 4	Cue 5	Cue 6	total
Total Häufigkeit	620	899	852	693	626	390	4080 ⁴
Erster Cue	202	435	216	244	246	97	1440

Nach dem ersten Cue wurde signifikant häufiger weitergesucht ($M = 47.0$, $SD = 7.4$), als dass die Informationssuche gestoppt wurde ($M = 13.0$, $SD = 7.4$; $t = 11.3$, $p < .001$). Insgesamt wurde in nur 22% aller 1440 Fälle nach einer Information gestoppt und demzufolge auch gemäss Success gestoppt. In 37% aller diskriminierenden Fälle wurde weitergesucht, in 63% gestoppt. Diskriminierte der Cue nicht, wurde in 98% der Fälle weitergesucht und in 2% gestoppt.

Die Hypothesen betreffend der Informationssuche nach zwei nicht-diskriminierenden Informationen lauten: Bei zwei negativ diskriminierenden Informationen, suchen die Personen bei Cues weiter, die höhere Ausprägungsraten haben und bei zwei positiven Cue-Ausprägungen, suchen die Personen häufiger bei Cues mit niedrigerer Ausprägungsrate weiter. Addiert man alle Cuekäufe, die nicht diskriminierten, dann ergibt dies 1290 Fälle, bei welchen der Cue mit zwei negativen Werten nicht diskriminierte und 1014 Fälle, bei welchen der Cue mit zwei positiven Werten nicht diskriminierte. Die darauf folgende Weitersuche wurde in zwei Kategorien klassifiziert: die Weitersuche bei einem Cue, der eine höhere Ausprägungsrate als der zuvor konsultierte Cue hat und die Weitersuche bei einem Cue, der eine niedrigere Ausprägungsrate hat. In Tabelle 8 wird ersichtlich, dass die Personen sich auf Gesamtebene genau so verhalten, wie in den Hypothesen postuliert wurde. Ein Chi-Quadrat-Test zeigt einen signifikanten Unterschied in dem Verhalten (Negativ: $p < .001$; Positiv: $p < .001$). Da die Weitersuche nach dem Aufdecken von Cue 1 und 6 immer unser Modell bestätigen, haben diese Fälle keine Vorhersagekraft. Schliesst man diese Fälle aus der Auswertung aus (siehe Tabelle 9), ist dennoch die gleiche Tendenz im Verhalten zu beobachten (Negativ, $p < .001$; Positiv: $p < .001$).

Tabelle 8: Weitersuche nach einer nichtdiskriminierender Information.

	Option A B	Cue hat eine höhere Ausprägungsrate	Cue hat eine niedrigere Ausprägungsrate	total
Nicht diskriminierend	- -	1006	284	1290
	+ +	123	891	1014
total		1129	1175	2304

Tabelle 9: Weitersuche nach einer nichtdiskriminierender Information, ohne Cue 1 und Cue 6.

	Option A B	Cue hat eine höhere Ausprägungsrate	Cue hat eine niedrigere Ausprägungsrate	total
Nicht diskriminierend	- -	691	283	974
	+ +	123	497	620
total		814	780	1594

Auf einer gesamten Betrachtungsebene gesehen, bestätigen diese Ergebnisse unsere Hypothesen im Bezug auf die weiterführende Informationssuche nach nicht-diskriminierender Information. Eine spezifische Analyse zeigt, dass bei der Weitersuche nach dem ersten Cue dieselbe Tendenz erhalten bleibt (siehe Tabelle 10) und die Unterschiede signifikant sind (Negativ: $p < .001$ und Positiv: $p < .001$).

⁴ Elf Personen haben nicht ihr gesamtes Budget ausgenutzt, fünf Personen haben ihr Budget überzogen. Das Programm lief weiter, um alle Daten zu erhalten, auch wenn die Versuchspersonen das Budget überzogen.

Tabelle 10: Weitersuche nach dem ersten nicht-diskriminierenden Cue.

	Option A B	Cue hat eine höhere Ausprägungsrate	Cue hat eine niedrigere Ausprägungsrate	total
Nicht diskriminierend	- -	369	135	504
	+ +	59	393	452
total		428	528	956

Das Verhalten nach dem ersten Cue kann danach kategorisiert werden, ob es modelkonform bzw. nicht modelkonform ist. Modelkonform bedeutet einerseits, dass die Leute bei einem Cue, der mit zwei negativen Informationen nicht diskriminiert, bei einem Cue weitersuchen, der eine höheren Ausprägungsrate hat, als der Cue selbst. Andererseits wenn der Cue mit zwei positiven Informationen nicht diskriminiert, suchen sie bei einem Cue mit niedrigerer Ausprägungsrate weiter. Alles andere Verhalten ist nicht Modelkonform.

Tabelle 11: Häufigkeiten, mit welchen der zweite Cue in der vorhergesagten Reihenfolge konsultiert wurde, d.h. ob diese Weitersuche nach dem ersten Cue modelkonform bzw. nicht modelkonform war.

	Cue 1	Cue 2	Cue 3	Cue 4	Cue 5	Cue 6	total
Modelkonform	165	206	62	76	172	81	762
Nicht Modelkonform	0	53	40	63	38	0	194

In Tabelle 11 wird nochmals deutlich, dass die Weitersuche nach Cue 1 und 6 immer unser Modell bestätigen und diese Fälle daher keine Vorhersagekraft haben. Auch wenn bei dem beobachteten Verhalten fast 80% (762 von 956 Fällen) des registrierten Verhaltens unser Modell bestätigen, sind die Daten so angelegt, dass in 76% aller Fälle (d.h. in 457 von 900 Fällen) das Modell passen müsste. Die beobachteten Daten unterscheiden sich in einem Chi-Quadrat Test zwar signifikant von den theoretisch erwarteten Werten, aber nur auf dem 5%-Niveau.

In der Nachbefragung zeigte sich, dass die Versuchspersonen die Ausprägungshäufigkeit, wenn explizit danach befragt, sehr gut angeben konnten. Im Mittel korrelierte die subjektive Rangreihe mit der objektiv erwarteten Rangreihe der sechs Cues sehr hoch ($r=.80$)⁵. Bewertet auf einer Skala von 1-6, ist den Personen das Symbol des Volleyball -Teams am sympathischsten ($M=4.4$, $SD=1.5$), das Symbol der Fahrradfabrik und der Linienbusse am zweit-sympathischsten ($M=4.0$, $SD=1.4$; $M=4.0$, $SD=1.3$); gefolgt vom Symbol des Krankenhauses ($M=3.3$, $SD=1.4$), dem Symbol des Parteibüros ($M=2.9$, $SD=1.7$) und zum Schluss dem Symbol der Tageszeitung ($M=2.8$, $SD=1.5$).

Diskussion

Es werden pro Runde mehr Informationen konsultiert, als dass ein *One Reason Decision Making* implizieren würde: die „Kaufhäufigkeit“ schwankt um die zwei bis drei Informationen pro Runde. Es wird nur in 22% aller Fälle gleich nach dem ersten Cue gestoppt, was die Hypothese eines Stoppens der Informationssuche gemäss Success in der Mehrzahl der Fälle falsifiziert. Dieses Verhalten macht jedoch insofern Sinn, als dass den Versuchspersonen für alle 60 Aufgaben 180 Informationen zur Verfügung standen, was aufgeteilt auf eine Runde um die drei Informationen ergibt. Dass gegen Ende der 60 Runden vermehrt Informationen

⁵ Werden die individuellen Korrelationen Fisher-Z transformiert, ergibt es sogar eine gemittelte Korrelation von $r=.90$.

gesucht wurden, ist ebenfalls durch diesen Hintergrund des fix vorgegebenen Informationsbudgets erklärbar. Eine Idee zur Verbesserung dieses methodischen Problems wäre ein flexibles Informationsbudget.

Wenn die Informationssuche frei wählbar ist, stellt sich die Frage welche Informationen die Personen bevorzugen und weshalb. Die Nützlichkeit, die ein Cue für eine Person hat, lässt sich anhand der gesamten Cuesuche, wie auch der Cuesuche des ersten Cues analysieren. Die Häufigkeit der Cuesuche zeigt, dass der Cue Parteibüro (Cue 2) insgesamt, aber auch als erster Cue präferiert wird. Die Präferenz dieses Cues bestätigt unsere Hypothese betreffend der Informationssuche des ersten Cues. Wir nahmen an, dass die Suche des ersten Cues gemäss Success erklärt werden kann. Nur stellt sich dabei die Frage, weshalb der Cue Volleyball-Team (Cue 4) – mit derselben Successrate – in seiner Nützlichkeit unterschätzt wird und viel seltener als erstes gewählt wird. Es gibt verschiedene Interpretationsmöglichkeiten für den Befund der Präferenz des Cue 2. Erstens könnte es sein, dass die Versuchspersonen die Cue-gütemasse nicht gut gelernt haben und deshalb eine verzerrte implizite Annahme über die Nützlichkeit dieses Cues machen. Womöglich überschätzen die Personen die Validität von Cue 2 und unterschätzen diese von Cue 4. Zweitens könnte es sein, da die Benennung der Cues nicht randomisiert wurde, dass das Parteibüro am nützlichsten erschien, um die Grösse einer chinesischen Stadt zu identifizieren. Dagegen spricht jedoch das Sympathierating, bei welchem das Parteibüro den zweitletzten Platz in der Nachbefragung belegte. Eine letzte, eher spekulative Vermutung ist, dass die Leute eine implizite Annahme über die Verteilung der Merkmale haben. Das heisst, da die natürliche Umwelt sich häufig aus J-shaped verteilten Merkmalen konstituiert, nehmen die Personen womöglich an, dass es viel mehr kleine Städte gibt, als grosse. Aus diesem Grund würden sie dann eher die Suche mit einem Cue beginnen, der bei kleinen Städten diskriminiert, als bei grossen Städten.

Die Hypothesen betreffend der abhängigen Informationssuche nach zwei nicht-diskriminierenden Informationen können in die prognostizierte Richtung bestätigt werden. Die Resultate zeigen, dass auf einem Gesamtniveau die Personen ihre weitere Informationssuche an die vorliegenden nicht-diskriminierende Information anpassen: bei zwei negativ diskriminierenden Informationen, suchen sie bei Cues weiter, die höhere Ausprägungsraten haben und somit eher bei kleinen Städten diskriminieren und bei zwei positiven Informationen suchen sie bei Cues weiter, die eine niedrigere Ausprägungsrate haben und eher bei grossen Städten diskriminieren. Somit verhalten sich die Personen signifikant häufiger modelkonform als nicht modelkonform. Wobei kritisch anzufügen ist, dass das Nullmodell in 76% der Fälle ein modelkonformes Verhalten postuliert. Zwei Variablen führen zu dieser Verteilung: Einerseits ergibt es eine ungleiche Verteilung von zwei positiven und zwei negativen Informationen, weil die Versuchspersonen die Informationen frei suchen können. Andererseits resultiert diese Ungleichverteilung aus der Anzahl der zur Verfügung stehenden Cues. Da dies sechs Cues waren, standen nach dem ersten Cue jeweils noch fünf Cues zur Verfügung, was ungleiche Erwartungswerte für die Cues mit tieferer und höherer Ausprägungsrate ergab. Deshalb wird in dem Experiment 2.1 der nicht diskriminierende Cue vorgegeben und die Anzahl der Cues von sechs auf fünf reduziert.

Die Ergebnisse zeigen eindeutig, dass die Personen ihr Suchverhalten an die vorhandene Information der Umwelt anpassen. Die streng sequentielle Suchregel von *Take The Best* trifft

nur in ganz wenigen Fällen zu: in den 202 Fällen, in welchen der valideste Cue Fahrradfabrik (Cue 6) als erstes konsultiert wurde, wurde die Weitersuche in 17 Fällen beim zweitvalidesten Cue Krankenhaus (Cue 5) fortgesetzt. Das bedeutet, dass die Informationssuche nur in 14% der Fälle beim validesten Cue begonnen wurde und in nur gerade 8% dieser Fälle bei dem zweitvalidesten Cue fortgesetzt wurde und somit nur ganz wenige Fälle der Suchregel von *Take The Best* entsprechen. Eventuell haben die Versuchspersonen aber auch einfach die Validität nicht adäquat gelernt und man müsste die subjektiven Rangreihen zu allen verschiedenen Gütemassen erheben und anhand derer die verschiedenen Suchstrategien identifizieren. Dagegen deuten die Resultate der Nachbefragung zum deklarativen Wissen der Ausprägungsrate darauf hin, dass die Personen diese Ausprägungsrate sehr gut explizit angeben können. Somit ist die nicht-diskriminierende Information nicht wertlos, sondern gibt den Personen einen Anhaltspunkt, in welchem Grössenbereich von Städten sie sich bewegen und dementsprechend passen sie ihre Informationssuche an. Es stellt sich nunmehr die Frage, ob die gefundenen Resultate von der Lernphase der Ausprägungsrate abhängig sind, oder ob dieser Effekt bei anderen Lernformaten ebenfalls zu finden ist.

Experiment 2.1: Ist die Anwendung der Ausprägungsrate abhängig vom Lernformat?

Anhand zweier Lerngruppen untersuchen wir den Einfluss des Lernformats auf die Anwendung der Ausprägungsrate. Die eine Lerngruppe A lernt analog zu den Experimenten von Bröder und Newell mit paarweisen Vorgaben der Cue-Informationen (Diskriminationslernen). Die andere Lerngruppe B lernt nur die Ausprägungsrate von den einzelnen Wahlalternativen (Ausprägungsrate-Lernen). Die Hypothese ist, dass die Lerngruppe B (Ausprägungsrate-Lernen) je nach Art der vorgegebener nicht-diskriminierender Informationen weitersuchen und so die Befunde aus Experiment 1 replizieren, die Lerngruppe A (Diskriminationslernen) jedoch nicht. Da im Experiment 1 ersichtlich wurde, dass unsere Annahmen mit der freien Informationssuche nicht exakt getestet werden können, wird in diesem Experiment der erste Cue vorgegeben und zwar ist dies immer der am höchsten diskriminierende Cue 3, der ein Hinweis für das Vorhandensein einer Tageszeitung ist.

Methode

Subjects

Es nahmen 60 Versuchspersonen teil, wobei die Lerngruppe A aus 17 Frauen und 13 Männern bestand, deren Durchschnittsalter 24 Jahre (Range: 18-45, SD=6.4) betrug. Darunter befanden sich 17 Studierende der Universität Zürich oder ETH. Das Durchschnittsalter der Lerngruppe B betrug 29 Jahre (Range: 17-49, SD = 10.0), wobei 16 Frauen und 14 Männer waren und 14 davon studierten.

Apparatus

Die Lernphase der Gruppe A, sowie die Testphase wurden in Microsoft Visual Basic 6.0 programmiert. Die Lernphase der Gruppe B wurde mit dem Powerpoint Programm erstellt. Das Experiment wurde vollständig auf einem PC mit Windows XP-Betriebssystem durchgeführt.

Stimuli, Design and Procedure

Das Experiment bestand jeweils aus einer Lernphase und einer Testphase, wobei ein In-between-subject Design mit zwei unterschiedlichen Lernphasen durchgeführt wurde. Die Aufgabe bestand in beiden Lernphasen darin, sich ein Gefühl der Nützlichkeit der Cues in Anbetracht auf die Beurteilung von Städtegrößen anzueignen. Die Lerngruppe A lernte die Nützlichkeit anhand von 60 paarweisen Vergleichen, die Lerngruppe B lernte diese Nützlichkeit anhand von 60 Städten analog zu Experiment 1. Bei den paarweisen Vergleichen wurden alle Cue-Ausprägungen von zwei Städten vorgegeben, die Personen mussten dabei jeweils auf die grössere der beiden Städte tippen und erhielten ein Feedback, ob sie mit ihrer Entscheidung richtig oder falsch lagen. Die Coverstory und das Information Board (siehe Abbildung 9) der Lerngruppe A sahen folgendermassen aus:

In diesem Experiment bist Du als Managerin bzw. Manager für eine internationale Firma tätig, die den lukrativen chinesischen Markt erobern will. Deine Aufgabe besteht darin, in 60 chinesischen Städten verschiedener Regionen jeweils eine Filiale zu eröffnen. Die PR-Abteilung Deiner Firma gibt Dir jeweils zwei in Frage kommende Städte pro Region vor. Was Du aber nicht weisst, ist, welche von den beiden Städten tatsächlich die grössere ist. Für die Gewinnmaximierung des Konzerns ist es aber von zentraler Bedeutung, dass Du die Filiale jeweils in der grösseren der beiden Städte eröffnest. Dadurch kommt der Konzern zu Mehreinnahmen, die indirekt wieder der örtlichen Bevölkerung zugute kommen. Deine Aufgabe besteht darin, zu entscheiden, welche der beiden Städte die grössere ist! Leider gibt die chinesische Regierung für ihre Städte keine Einwohnerzahlen bekannt. Es stehen jedoch indirekte Hinweise zur Verfügung, die Dir helfen, die grössere der beiden Städte zu identifizieren. Das Vorhandensein von Fahrradfabrik, Krankenhaus, Linienbusse, Parteibüro, Tageszeitung und Volleyball-Team macht mit einer gewissen Wahrscheinlichkeit eine richtige Aussage über die Grösse der Städte. Im ersten Teil geht es nun darum, Dich anhand aller Hinweise für die grössere der beiden Städte zu entscheiden und Dir ein Gefühl für die Nützlichkeit der einzelnen Hinweise zu bilden. Die Zeit, die Du für Deine Entscheidungen benötigst, spielt keine Rolle.

Welche Stadt hat mehr Einwohner?

Wenhai Xipong

Antwort: Wenhai OK

FALSCH

Wenhai	Hinweis	Xipong
-	Volleyball-Team	-
-	Parteibüro	+
-	Krankenhaus	-
+	Linienbusse	-
-	Tageszeitung	+

Weiter

Abbildung 9: Lernphase der Lerngruppe A

In der Testphase zur weiterführenden Suche wurde jeweils ein Cue vorgegeben und zwar derjenige mit der höchsten Diskriminationsrate (Cue 3). Die Position der Cues wurde in jedem Durchgang zufällig geändert, ansonsten war das Information Board analog zur Testphase des Experimentes 1 aufgebaut. Die Testphase umfasste insgesamt 36 Runden, wobei den Versuchspersonen 72 Informationen zur Verfügung standen. Die Instruktion lautete folgendermassen:

Deine Firma freut sich, Dich als ManagerIn weiterbeschäftigen zu können. Sie beauftragt Dich, neue Filialen in weiteren chinesischen Städten verschiedener Regionen zu eröffnen. Die zuständige Abteilung Deiner Firma gibt Dir jeweils wieder zwei in Frage kommende Städte pro Region vor. Für die Gewinnmaximierung des Konzerns ist es von zentraler Bedeutung, dass Du die Filiale jeweils in der grösseren der beiden Städte eröffnest.

Leider weisst Du diesmal nicht, ob eine Fahrradfabrik, ein Krankenhaus, Linienbusse, ein Parteibüro, eine Tageszeitung oder ein Volleyball-Team bei diesen Städten vorhanden sind. Du kannst jedoch eine chinesische Beamtin weiterhin um Hinweise bitten und dir ein Bild über die Grösse der zur Auswahl stehenden Städte machen. In den nächsten vier Aufträgen hast du einmal mehr und einmal weniger Hinweise zur Verfügung. Deine Aufgabe besteht jedoch immer darin, auszuwählen, welche der beiden Städte die grössere ist!

Für die Eröffnung der folgenden 36 Filialen gilt:

Der erste Hinweis steht dir jeweils uneingeschränkt zur Verfügung. Grundsätzlich kannst du pro Filiale so viele Hinweise wie gewünscht erfragen, mit einer Einschränkung: Für die 36 Entscheide stehen dir - neben dem ersten Hinweis - insgesamt 72 Hinweise zur Verfügung.

Teile die beschränkte Zahl an Hinweisen so ein, dass du in allen Aufgaben genügend Hinweise aufdecken kannst und am Ende nicht raten musst.

Deine Aufgabe besteht immer noch darin, auszuwählen, welche der beiden Städte die grössere ist!

Die Datenlage basierte auf folgenden Cue-Gütemassen, wobei die Städtegrössen auch wieder linear verteilt waren:

Tabelle 12: Validität, Diskriminationsrate, Ausprägungsrate, Success und Usefulness der fünf Cues.

	Validität (V)	Diskriminationsrate (DR)	Ausprägungsrate (AR)	Success	Usefulness
Cue 1: Linienbusse	.75	.20	.90	.55	.15
Cue 2: Parteibüro	.85	.43	.70	.65	.37
Cue 3: Tageszeitung	.70	.50	.50	.60	.35
Cue 4: Volleyball-Team	.92	.43	.30	.68	.40
Cue 5: Krankenhaus	1.00	.20	.10	.60	.20

Resultate

Von den 36 Runden wurden nur die ersten 30 Runden ausgewertet, da in Experiment 1 eine Verzerrung der Informationssuche gegen Ende des Experimentes aufgrund des beschränkten Kostenbudgets gefunden wurde. In Tabelle 13 sind sowohl die Häufigkeiten der gesamten Informationssuche, wie auch des ersten konsultierten Cues dargestellt.

Tabelle 13: Prozentuale Häufigkeit der gesamten Informationssuche und ersten Cue bei Lerngruppe A (Diskriminationslernen) und Lerngruppe B (Ausprägungsrate-Lernen), wobei der Cue 3 jeweils aufgedeckt vorgegeben wurde.

Gesamte Häufigkeit	Cue 1	Cue 2	Cue 3	Cue 4	Cue 5
Lerngruppe A	.22	.24	0	.29	.25
Lerngruppe B	.21	.33	0	.26	.20
Erster Cue					
Lerngruppe A	.17	.18	0	.40	.26
Lerngruppe B	.16	.41	0	.23	.19

Die beiden Lerngruppen präferierten nach dem Cue 3 unterschiedliche Cues: die Lerngruppe A präferierte den Cue 4 stärker als die Lerngruppe A ($t=2.15$, $p<.05$), die Lerngruppe B hingegen konsultierte den Cue 2 häufiger als die Lerngruppe A ($t=-2.97$, $p<.01$). Über die gesamte Cue-Kaufhäufigkeit löst sich dieser Unterschied wieder auf (Cue 2: $t=1.8$, $p=.07$; Cue 4: $t=.07$; $p=.48$).

Bei beiden Lerngruppen wurde häufiger nach der ersten vorgegebenen Information weitergesucht als gestoppt, bei der Lerngruppe A wurde in 87% aller Fälle und der Lerngruppe B in 83% aller Fälle weitergesucht und somit die Stoppregel von Success verletzt. Wenn weitergesucht wurde, dann geschah dies in 57% bzw. 60% der Fälle nach einer nicht-diskriminierenden Information, und in 43% bzw. 40% der Fälle nach einer diskriminierenden Information. Das Weitersuchen nach einer diskriminierenden Information verstösst gegen die Stoppregel von *Take The Best*.

In Tabelle 14 und 15 wird ersichtlich, dass die beiden Lerngruppen sich bei der Weitersuche unterschiedlich verhielten. Obwohl beide Gruppen sich in etwa gleich häufig modelkonform verhielten ($t=-0.7$, $p=.48$), d.h. sie suchten nach zwei negativen Informationen bei einem Cue mit höheren Ausprägungsrate weiter und nach zwei positiven Informationen bei einem Cue mit niedrigerer Ausprägungsrate, unterschieden sie sich betreffend unseren Vorhersagen der abhängigen Suchregel.

Tabelle 14: Weitersuche nach dem vorgegebenen Cue 3 bei der Lerngruppe A (Diskriminationslernen).

	Option A B	Cue hat eine höhere Ausprägungsrate	Cue hat eine niedrigere Ausprägungsrate	total
nicht diskriminierend	- -	92	146	238
	+ +	54	155	209
total		146	301	447

Tabelle 15: Weitersuche nach dem vorgegebenen Cue 3 bei der Lerngruppe B (Ausprägungsrate-Lernen).

	Option A B	Cue hat eine höhere Ausprägungsrate	Cue hat eine niedrigere Ausprägungsrate	total
nicht diskriminierend	- -	146	88	234
	+ +	101	112	213
total		247	200	447

Die Lerngruppe A verhielt sich in 55% (siehe Tabelle 14: 247 von 447 Fällen) der Fälle modelkonform, was im Binomialtest signifikant wird ($p<.05$). Der Effekt kommt aber hauptsächlich durch das Verhalten nach zwei positiven Informationen zu Stande, denn die Versuchspersonen verhielten sich häufiger ($t=-2.2$, $p<.05$) nach zwei positiven Informationen modelkonform (im Mittel in 33% der Fälle), als nach zwei negativen Informationen (in 22% der Fälle). Der Haupteffekt wird vor allem durch fünf Personen (von 30) verursacht, die sich in mehr als 70% der Fälle modelkonform verhielten.

Die Lerngruppe B suchte nach dem nicht-diskriminierenden Cue in 58% (siehe Tabelle 15: 258 von 447) der Fälle, modelkonform weiter. Das heisst die Versuchspersonen suchten häufiger modelkonform weiter, als nicht modelkonform (Binomialtest mit $p<.01$). Im Gegensatz zur Lerngruppe A, suchte die Lerngruppe B häufiger bei zwei negativen Informationen modelkonform weiter (33%), als bei den positiven Informationen (25% der Fälle). Dieser Effekt kommt hauptsächlich durch neun Personen zustande (die sich in mehr als 70% der Fälle modelkonform verhielten).

Beide Lerngruppen suchten in Abhängigkeit der nicht-diskriminierenden Information bei Cues weiter, welche eine entsprechend höhere oder tiefere Ausprägungsrate hatten. Dennoch zeigen sich Unterschiede zwischen den beiden Gruppen: die Lerngruppe A präferierte insgesamt überzufällig häufig einen Cue mit niedrigerer Ausprägungsrate (vs. dem Zufall von .5; One Sample T-Test: $t=3.0$, $p<.01$), die Lerngruppe B hingegen suchte insgesamt gleich häufig

bei einem Cue mit hoher Ausprägungsrate, wie bei einem Cue mit niedriger Ausprägungsrate (One Sample T-Test: $t=1.0$, $p=.34$) weiter. Es stellt sich somit die Frage, ob die Versuchspersonen in den beiden Lernbedingungen unterschiedliche Cues präferierten, weil sie deren Güte unterschiedlich gelernt hatten.

Experiment 2.2: Ist das Lernen der Cue- Gütemasse Validität und Diskriminationsrate abhängig vom Lernformat?

Anhand derselben beiden Lerngruppen untersuchten wir deshalb die Lernbarkeit der Gütemasse Validität, Diskriminationsrate und deren Anwendung beim Success. Wir nehmen an, dass die Lerngruppe A (Diskriminationslernen) sowohl die Validitäten, wie auch die Diskriminationsraten gut lernen sollte, wobei die Lerngruppe B die Validitäten etwas weniger gut lernen sollte, die Diskriminationsrate aber ebenfalls über die Ausprägungsrate erschliessen können sollte. Getestet wird das Wissen über die Gütemasse anhand von Verhaltenstests analog zum Äffchenexperiment. In drei unterschiedlichen Bedingungen mussten die Personen die Reihenfolge der Cues an die Fragestellung des V) validesten Cues, D) besten diskriminierenden Cue, S) success-vollsten Cue anpassen. Einerseits sagt die Häufigkeit des ersten konsultierten Cues etwas über die gelernte Nützlichkeit der Cues aus, andererseits aber auch die Häufigkeit, mit welcher der erste Cue in den 12 Runden der jeweiligen Bedingung immer derselbe war.

Methode

Subjects und Apparatus

Dieselben Versuchspersonen, wie in Experiment 2.1 beschrieben, führten diesen zweiten Teil durch.

Stimuli, Design and Procedure

Im Anschluss an Experiment 2.1 wurden eine zweite Testphase und eine Nachbefragung zur Ausprägungsrate angehängt. In der Nachbefragung mussten die Versuchspersonen die fünf Cues gemäss der Ausprägungsrate in eine Rangreihe bringen. Diese Testphase bestand aus drei randomisierten Bedingungen, in welchen die Validität (V), die Diskriminationsrate (D) und Success (S) abgefragt wurde. Im Experiment sahen die Instruktionen für die drei Bedingungen folgendermassen aus:

- V) Für die Eröffnung der folgenden 12 Filialen gilt:
Du darfst grundsätzlich so lange Hinweise erfragen, bis du einen Hinweis erhältst, der bei den zwei Städten unterschiedliche Informationen zeigt. Danach musst du dich entscheiden.

Wähle deshalb die Reihenfolge der Hinweise gut aus (denn sie sind zufällig angeordnet) und eröffne möglichst viele der 12 Filialen in der grösseren Stadt!

Deine Aufgabe besteht immer noch darin, auszuwählen, welche der beiden Städte die grössere ist!
- D) Für die Eröffnung der folgenden 12 Filialen gilt:

Du kannst grundsätzlich alle Hinweise pro Filiale erfragen, solange der erste aufgedeckte Hinweis bei beiden Städten unterschiedliche Information zeigt. Wenn aber der erste Hinweis bei beiden Städten dieselbe Information liefert, musst du dich sofort entscheiden und die grössere Stadt erraten.

Wähle die Reihenfolge der Hinweise gut aus, denn sie sind zufällig angeordnet!

Deine Aufgabe besteht immer noch darin, auszuwählen, welche der beiden Städte die grössere ist!

S) Für die Eröffnung der folgenden 12 Filialen gilt:

Du kannst nur noch einen einzigen Hinweis pro Filiale erfragen. Deine Entscheidung wird jeweils auf dem ersten aufgedeckten Hinweis beruhen. Falls der von dir gewählte Hinweis bei beiden Städten dieselbe Information zeigt, kannst du nur noch raten, welches wohl die grössere Stadt ist.

Wähle also den Hinweis gut aus.

Deine Aufgabe besteht immer noch darin, auszuwählen, welche der beiden Städte die grössere ist!

Resultate

Die Lernbarkeit der Gütemasse Validität und Diskriminationsrate und deren Kombination im Success sollten sich zwischen den beiden Lerngruppen unterscheiden. Wir nehmen an, dass die Lerngruppe A (Diskriminationslernen) sowohl die Validitäten, wie auch die Diskriminationsraten gut lernen. Die Lerngruppe B sollte die Validitäten weniger gut lernen, die Diskriminationsrate aber genauso gut, weil sie diese über die Ausprägungsrate erschliessen können. Wie gut diese Gütemasse gelernt wurden, wurde in dieser Auswertung über das Verhalten getestet. Einerseits wurde die Häufigkeit des ersten konsultierten Cues (siehe Tabelle 16), andererseits aber auch die Konsistenz dieser Konsultation, d.h. die Häufigkeit, mit welcher die Cues in den 12 Runden der jeweiligen Bedingung konsultiert wurden, ausgewertet.

Tabelle 16: Häufigkeit der Konsultation des ersten Cues in den drei Bedingungen Validität, Diskriminationsrate und Success bei den beiden Lerngruppen A (Diskriminationslernen) und Lerngruppe B (Ausprägungsrate-Lernen), wobei jede Bedingung aus 12 Durchgängen bestand.

	Cue 1	Cue 2	Cue 3	Cue 4	Cue 5	% Lerner
Lerngruppe A: Validität	.13	.10	.07	.30	.40	.60
Lerngruppe B: Validität	.15	.21	.15	.20	.30	.33
Lerngruppe A: Diskriminationsrate	.09	.18	.32	.28	.14	.30
Lerngruppe B: Diskriminationsrate	.06	.39	.25	.16	.14	.50
Lerngruppe A: Success	.06	.26	.15	.36	.16	.37
Lerngruppe B: Success	.10	.37	.15	.28	.10	.47

Bei der Bedingung zur Testung der gelernten Validität, konsultierten beide Lerngruppen den validesten Cue (Cue 5) am häufigsten als erstes (Chi-Test der Lerngruppe A: $p < .001$, Chi-Test der Lerngruppe B: $p < .001$). Die Lerngruppe B hat den validesten Cue etwas weniger häufig angeschaut als die Lerngruppe A, aber nicht signifikant ($t = 1.0$, $p = .31$). In der Testbedingung zur Diskriminationsrate wurde der am stärksten diskriminierendste Cue (Cue 3) von der Lerngruppe A am häufigsten als erstes angeschaut (Chi-Test: $p < .001$). Die Lerngruppe B präferiert hingegen den Cue 2 in dieser Bedingung (Chi-Test: $p < .001$), gefolgt von Cue 3 (Chi-Test: $p < .02$). Die Häufigkeit der Konsultation von Cue 3 unterscheidet sich zwischen den beiden Lerngruppen nicht signifikant ($t = 0.9$, $p = .4$), dafür aber die Häufigkeit von Cue 2 ($t = -3.3$, $p < .01$). Müssen die Versuchspersonen die beiden Cue-Gütemasse verrechnen und gemäss Success anwenden, dann bevorzugt die Lerngruppe A den successvollsten Cue 4 (Chi-Test: $p < .001$), die Lerngruppe B wiederum den zweit-successvollsten Cue 2 (Chi-Test: $p < .001$). Die Unterschiede sind jedoch nicht signifikant (Cue 2: $t = -1.2$, $p = .20$ und Cue 4: $t = 1.7$, $p = .38$).

Wie konsistent war das Informationssuchverhalten der einzelnen Versuchspersonen beim ersten Cue über die zwölf Runden hinweg? Als konsistent wird ein Verhalten dann kategorisiert,

wenn in mehr als 75 % der Fälle derselbe Cue als erstes angeschaut wurde. In der Lerngruppe A wählten 19 Versuchspersonen bei der Validitätsbedingung konsequent denselben Cue, bei der Diskriminationsrate neun Personen und bei der Success-Bedingung zehn Personen. Bei der Lerngruppe B wählten 14 Versuchspersonen bei der Validitätsbedingung konsequent denselben Cue, zehn Personen bei der Bedingung der Diskriminationsrate und elf Personen bei der Success-Bedingung.

Möchte man die Versuchspersonen danach analysieren, ob sie die Gütemasse tatsächlich gelernt haben, dann kann ein hypothetisches Lernkriterium angenommen werden. Ein niedrig angesetztes Lernkriterium wäre dann erreicht, wenn eine Versuchsperson den ersten und/oder zweiten je nach Bedingung nützlichster Cue, in mehr als 10 der 12 Runden auswählt. In der letzten Spalte in Tabelle 16 ist die Anzahl Versuchspersonen aufgeführt, die diesem Kriterium entsprechen. Bei der Lerngruppe A lernen die meisten Personen die Validität am besten, bei der Lerngruppe B die Diskriminationsrate und den Success.

In der Nachbefragung korrelierte die subjektive Ausprägungsrate der Versuchspersonen der Lerngruppe A mit $r = .39$, wobei 12 Personen die Rangreihe nahezu perfekt einschätzten ($r > .70$). Bei der Lerngruppe B korrelierte die Ausprägungsrate im Mittel mit $r = .65$, wobei 19 Personen die Ausprägungsrate fast perfekt replizieren konnten. Die Rangkorrelationen der beiden Versuchspersonengruppen unterscheiden sich signifikant ($t = -2.1$, $p < .05$).

Diejenigen zwölf Personen, welche in Experiment 2.1 adaptiv auf die Ausprägung der nicht-diskriminierenden Information reagierten, d.h. sich modelkonform verhielten, deren Korrelationen waren jedoch nicht höher ($r = .58$) als die Korrelationen aller Versuchspersonen.

Diskussion Experiment 2.1 und 2.2

In dem ersten Experiment suchen die beiden Lerngruppen überzufällig häufig weiter, auch nachdem sie schon eine diskriminierende Information erhalten haben, und verletzen somit die Stoppregel von Success und *Take The Best*. Dies ist jedoch nicht weiter erstaunlich, da den Versuchspersonen ein Informationsbudget von durchschnittlich zwei Cues pro Runde zur Verfügung stand. Wenn die Personen dann weitersuchen, verhalten sich beide Lerngruppen modelkonform. Jedoch zeigt sich nur bei 20% aller Personen, also bei 12 von 60 eine grundlegende Adaptivität an die nicht-diskriminierende Information.

Die beiden Lerngruppen suchen nach dem vorgegebenen Cue der Tageszeitung bei verschiedenen Cues weiter. Über die ganze Informationssuche präferiert die Lerngruppe A den Cue 4 (Volleyball-Team), die Lerngruppe B den Cue 2 (Parteibüro). Gleich nach dem Tageszeitungs-Cue schaut die Lerngruppe A häufig einen Cue mit niedrigerer Ausprägung an (z.B. Cue 4), was den Haupteffekt erklärt. Es kann vermutet werden, dass diese Lerngruppe ihre Informationssuche nach der Validität richtet, da die Cues mit den niedrigeren Ausprägungsraten gleichzeitig die Cues sind, welche die höheren Validitäten haben. Die Lerngruppe B hingegen sucht folgend auf den Tageszeitungs-Cue gleich häufig bei Cues mit niedriger oder höherer Ausprägungsrate weiter. Das bedeutet, dass unsere Hypothese betreffend der abhängigen Suchregel nur für die Lerngruppe B, also für die Lerngruppe, welche das Ausprägungsrate-Lernen absolvierte, bestätigt werden kann und untermauert damit unsere anfängliche Hypothese betreffend der Anwendung der Ausprägungsrate in Abhängigkeit vom Lernformat.

Zudem zeigen die Daten aus der Nachbefragung, dass die Lerngruppe A, welche die Cue-Gütemasse anhand des klassischen paarweisen Settings erwerben, die Ausprägungsrate weniger gut lernen als die Lerngruppe B.

Es drängt sich somit die Frage auf, ob die beiden Lerngruppen nicht nur bei der Ausprägungsrate, sondern auch bei den Gütemassen der Validität und Diskriminationsrate unterschiedliches gelernt haben. Die Auswertungen anhand eines angenommenen Lernkriteriums zeigen durchgehend schwache Lernerfolge und dies obwohl das Kriterium sogar noch eher niedrig angesetzt wurde: nur knapp 50% aller Personen lernen das eine oder andere Gütemasse so, dass sie es richtig anwenden können. Vergleicht man die beiden Lerngruppen, dann kann unsere Hypothese, dass die Lerngruppe A die Validität etwas besser lernt, angenommen werden. Analysiert man das Suchverhalten des ersten Cues, dann haben beide Gruppen die Validität zwar etwa gleich gut gelernt, aber es können bei der Lerngruppe A eindeutig mehr Personen als Validitätslerner klassifiziert werden. Die Hypothese, dass beide Gruppen die Diskriminationsrate etwa gleich gut lernen, muss jedoch falsifiziert werden. Beim Lernen der Diskriminationsrate sind die beiden Lerngruppen im Mittel zwar etwa gleich gut, aber zählt man auch hier die klassifizierten Lerner aus, ist die Lerngruppe B eindeutig besser. Bei der Kombination von V*D gemäss Success, sucht die Lerngruppe A den successvollsten Cue am häufigsten, aber sowohl auf der Mittelwertebene, wie auch den klassifizierten Lernern findet sich kein Unterschied zwischen den beiden Lerngruppen. Die in Experiment 2.1 gefunden Präferenz von Cue 4 bei der Lerngruppe A und Cue 2 bei der Lerngruppe B wird aber in der Success Bedingung repliziert. Das bedeutet, dass die beiden Lerngruppen zwei unterschiedliche Cues als nützlichsten Cue wahrnehmen. Es kann vermutet werden, dass die Lerngruppe A die Validität des Cue 4 überschätzt, die Lerngruppe B hingegen die Diskriminationsrate des Cue 2. Bei der Lerngruppe A bestimmt diese Wahrnehmung die Weitersuche nach nicht-diskriminierender Information, da sie signifikant häufiger bei einem Cue mit tiefer Ausprägungsrate, d.h. mit hoher Validität weitersucht. Bei der Lerngruppe B beeinflusst diese Wahrnehmung die Weitersuche vermutlich ebenfalls, aber nicht in demselben starken Ausmass. Diese Lerngruppe sucht unabhängig von der generellen Präferenz von Cue 2, bei einem Cue mit höherer bzw. tieferer Ausprägungsrate adaptiv an die Informationslage weiter. Diejenigen Personen, die sich in Experiment 2.1 bei der Weitersuche adaptiv verhalten, sind deshalb auch hauptsächlich Personen aus dieser Lerngruppe B.

Wir nahmen an, dass die Konsistenz mit welcher die Cues in den jeweiligen Bedingungen konsultiert werden, ein weiteres Indiz dafür ist, wie gut ein Gütemass gelernt wurde. Gemäss den Auswertungen hätten demnach die meisten Personen die Validität am besten gelernt. Es gibt jedoch eine nicht auszuschliessende Alternativerklärung für dieses Verhalten. Es könnte sein, dass bei den Testbedingungen zu Success und Diskriminationsrate die Personen schnell merken, ob sie jetzt mit dem jeweils „richtigen“ Cue angefangen haben oder nicht (denn wenn nicht, dürfen sie nicht weitersuchen und müssen raten), was bei der Validität nicht der Fall ist. So ändern die Personen bei den Bedingungen der Diskriminationsrate und Success vermutlich schneller ihre Suchstrategie aufgrund von negativem Feedback, als bei der Validitätsbedingung, wo sie dieses Feedback viel weniger schnell erhalten.

Im Grossen und Ganzen bestätigen diese Befunde unserer Hypothesen. Die Lerngruppe, welche anhand der Paarvergleiche lernt, lernt die Validität besser, als die Lerngruppe, welche

anhand der Ausprägungsrate lernt, jedoch nur marginal. Hingegen lernen diejenigen mit der Ausprägungsrate-Lernphase einerseits die Ausprägungsrate besser und können anhand derer auch relativ gut auf die Diskriminationsrate schliessen und diese sogar noch etwas besser anwenden als die andere Gruppe. Jedoch zeigen sich auch in diesen Auswertungen grosse individuelle Unterschiede. Fast die Hälfte der Versuchspersonen hat Mühe mit dem Erlernen der Gütemasse, unabhängig von dem Lernformat. Eventuell müssten mehr Lerndurchgänge durchgeführt werden. Da jedoch die Lernphase des Ausprägungslernen den Versuchspersonen schon zu lange war, müsste man diese aktiver gestalten. Erstens berichteten die VersuchsteilnehmerInnen, dass die Versuchspersonen in dieser Bedingung, obwohl sie nicht länger dauerte als das Diskriminationslernen, gelangweilt waren und nach dem Sinn der Aufgabe fragten. Zweitens sollten die beiden Lernbedingungen in der Aufgabenstellung ähnlicher gestaltet werden.

Dass die Ausprägungsrate-Lerner die Diskriminationsrate besser lernen als die andere Gruppe, ist doch etwas erstaunlich. Man könnte vermuten, dass das Lernen der Diskriminationsrate durch das Testen der abhängigen Suchregel kontaminiert wurde. In Experiment 2.2 wurde das gelernte Wissen über die Cue-Gütemasse nämlich immer erst nach der Anwendung der Ausprägungsrate, mit jeweils dem vorgegebenen stark diskriminierenden Cue, getestet. Da eine Interaktion zwischen dieser Testphase der Ausprägungsrate und dem Wissenstest über die Diskriminationsrate möglich wäre, soll in einem dritten Experiment nur das Cue-Gütewissen getestet werden.

Experiment 3: Implizites versus explizites Lernen der Cue-Gütemasse

Unsere Hauptfragestellung lautet erneut: Unter welcher Lernbedingung lernen die Personen die Cue-Gütemasse Validität, Diskriminationsrate und Success besser? Zudem soll untersucht werden, ob ein Unterschied zwischen dem impliziten und dem expliziten Wissen dieser Gütemasse besteht. Indem das implizite Wissen durch das Verhalten im Experiment und das explizite Wissen durch eine Nachbefragung im Fragebogen getestet wird, soll eine mögliche Differenz dieses Wissens explorativ eruiert werden.

Methode

Stichprobe

24 Versuchspersonen nahmen im Rahmen eines Experimentalpsychologischen Praktikums teil. Die Lerngruppe A bestand aus 5 Frauen und 7 Männern, deren Durchschnittsalter betrug 21.9 Jahre (Range: 19-27, SD= 2.3), darunter befanden sich 6 Studierende der Universität Zürich oder ETH. Das Durchschnittsalter der Lerngruppe B betrug 31.6 Jahre (Range: 21-57, SD = 14.6), wobei 6 Frauen und 6 Männer darunter waren und 9 davon studierten. Die beiden Stichproben unterscheiden sich signifikant in ihrem Alter ($t=2.4$, $p<.05$).

Apparatus

Die Lernphase der Gruppe A, sowie die Testphase wurden in Microsoft Visual Basic 6.0 programmiert. Die Lernphase der Gruppe B wurde mit dem Powerpoint Programm erstellt. Das Experiment wurde vollständig auf einem PC mit Windows XP-Betriebssystem durchgeführt, die Nachbefragung mit einem Fragebogen (siehe Anhang) erhoben.

Stimuli, Design and Procedure

Das Experiment bestand wiederum aus einer Lernphase und einer Testphase mit den drei Bedingungen, wobei ein In-between-subject Design mit zwei unterschiedlichen Lernphasen, analog zu dem vorherigen Experiment, durchgeführt wurde. Die Lerngruppe A lernte die Nützlichkeit der Cues anhand von 60 paarweisen Vergleichen (Diskriminationslernen) und die Lerngruppe B lernte diese Nützlichkeit anhand von 60 Städten anhand einzelner Städte (Ausprägungsrate-Lernen). Im Unterschied zu Experiment 2 wurden gleich nach der Lernphase die drei Bedingungen zur Validität, Diskriminationsrate und Success randomisiert angeschlossen und das Experiment mit einem Fragebogen zum expliziten Wissen der Gütemasse abgeschlossen. Die Datenlage und die Verteilung der Städtegrößen entsprachen denjenigen aus Experiment 2.

Resultate

Das Wissen über die Cue-Gütemasse wird implizit und explizit getestet. Das implizite Wissen wird anhand der Häufigkeit, mit welcher die Cues mit der entsprechend besten Cue-Güte in den jeweiligen Bedingungen auch als erstes konsultiert werden, ausgewertet. In Tabelle 17 sind die Häufigkeiten des ersten Cues in den drei Bedingungen prozentual dargestellt. Das explizite Wissen wird anhand der Korrelation zwischen den objektiven Rangreihen und den subjektiven Rangreihen der Gütemasse ausgewertet (siehe Tabelle 18).

Tabelle 17: Häufigkeit der Konsultation des ersten Cues in den drei Testphasen V (Validität), D (Diskriminationsrate) und S (Success) bei den beiden Lerngruppen A (Diskriminationslernen) und B (Ausprägungsrate-Lernen).

	Cue 1	Cue 2	Cue 3	Cue 4	Cue 5	% Lerner
Lerngruppe A: Validität	.12	.21	.15	.20	.33	.33
Lerngruppe B: Validität	.25	.13	.13	.09	.41	.25
Lerngruppe A: Diskriminationsrate	.13	.19	.15	.29	.24	.25
Lerngruppe B: Diskriminationsrate	.14	.28	.25	.18	.15	.42
Lerngruppe A: Success	.12	.28	.13	.26	.22	.17
Lerngruppe B: Success	.21	.26	.16	.15	.23	.08

Bei der Bedingung der Validität wird bei der Lerngruppe A keiner der Cues signifikant häufiger konsultiert (z.B. Cue5 vs. Cue 1: $t=2.0$, $p=.07$). Bei der Lerngruppe B wird der Cue 5 signifikant häufiger als der Cue 2, Cue 3 und Cue 4 gekauft (jeweils $p<.05$). Bei der Bedingung der Diskriminationsrate konsultiert die Lerngruppe A keinen Cue häufiger als die anderen (z.B. Cue 4 vs. Cue 1: $t=1.4$, $p=.19$), die Lerngruppe B auch nicht (z.B. Cue 2 vs. Cue 1: $t=1.2$, $p=.25$). In der Anwendung von Success konsultierte die Lerngruppe A ebenfalls keinen Cue häufiger (z.B. Cue 2 vs. Cue 1: $t=1.6$, $p=.13$), die Lerngruppe B auch nicht (Cue 2 vs. Cue 4: $t=1.1$, $p=.30$). Das bedeutet, dass einzig die Lerngruppe B das Gütemass der Validität im Verhaltenstest richtig anwendet.

Eine Möglichkeit das gelernte implizite Wissen über die Gütemasse auf individueller Ebene zu eruieren, ist die Anzahl Versuchspersonen auszuzählen, die einem bestimmten Kriterium entsprechen, z.B. eine Person hat den ersten und/oder zweiten infrage kommenden Cue, je nach Bedingung, in mehr als 10 der 12 Fälle ausgewählt. Bei der Lerngruppe A lernte ein Drittel die Validität am besten, bei der Lerngruppe B lernte knapp die Hälfte die Diskriminationsrate am besten. Diese Befunde replizieren diejenigen aus dem Experiment 2.

Um das explizite Wissen zu testen, wurde die Rangkorrelation zwischen den objektiven Cue-Gütemasse und den subjektiven Angaben aus der Nachbefragung gerechnet (siehe Tabelle 18).

Tabelle 18: Rangkorrelationen der subjektiven Rangreihe im Fragebogen mit der objektiven Rangreihe der drei Bedingungen V (Validität), D (Diskriminationsrate) und S (Success) bei den beiden Lerngruppen A (Diskriminationslernen) und Lerngruppe B (Ausprägungsrate-Lernen).

	Lerngruppe	Validität (V)	Diskriminationsrate (DR)	Ausprägungsrate (AR)	Success
	A	0.9	-0.1	0.5	-0.5
	A	0.8	0.5	-0.3	1.0
	A	0.7	0.9	0.7	0.8
	A	0.7	0.1	0.8	0.4
	A	0.7	-0.5	0.3	-0.4
	A	0.5	0.1	-0.4	0.2
	A	0.5	-0.2	0.2	-0.2
	A	0.3	0.9	-0.3	-0.6
	A	0.3	-0.5	0.4	-0.8
	A	0.2	0.2	0.7	0.5
	A	0.2	-0.4	-0.2	-0.2
	A	-0.1	0.7	0.5	0.7
	B	0.9	-0.6	0.9	0.4
	B	0.9	-0.6	1.0	0.4
	B	0.6	0.2	1.0	0.7
	B	0.3	0.0	1.0	-0.7
	B	0.3	-0.2	1.0	0.6
	B	-0.1	0.6	0.9	0.7
	B	-0.1	0.2	0.9	0.6
	B	-0.3	0.2	0.6	-0.7
	B	-0.4	0.5	0.8	-0.6
	B	-0.5	0.7	0.9	0.6
	B	-0.5	0.0	0.9	-0.7
	B	-0.6	0.3	0.1	-0.7
Mittelwert	A	0.48	0.15	0.24	0.08
Mittelwert (Fisher Z)	A	0.54	0.27	0.29	0.18
Mittelwert	B	0.04	0.10	0.83	0.04
Mittelwert (Fisher Z)	B	0.13	0.11	0.82⁶	0.03

Wir gehen davon aus, dass die Versuchspersonen die Gütemasse gelernt haben, wenn ihre subjektive Rangreihe mit der objektiven Rangreihe mit mindestens einer Korrelation von .7 übereinstimmt (in Tabelle 18 fett hervorgehoben). Bei der Lerngruppe A sind dies fünf Personen, die die Validität der Cues in die richtige Reihenfolge bringen; drei Personen, die die Diskriminationsrate adäquat einschätzen können; drei Personen, die die Ausprägungsrate richtig beurteilen und drei Personen, die die Gütemasse gemäss Success kombinieren können. Vier Personen scheinen keines der Gütemasse richtig beurteilen zu können. Bei der Lerngruppe B können zwei Personen die Validität adäquat einschätzen, eine Person die Diskriminationsrate; elf Personen die Ausprägungsrate und zwei Personen, die Cues gemäss Success richtig ordnen. Zwei Personen können keine der Cue-Gütemasse richtig beurteilen. Die beiden Gruppen unterscheiden sich in den Rangkorrelationen weder bei der Einschätzung von der Validität ($t(22) = 0.45$, $p = .65$), noch bei der Diskriminationsrate ($t(22) = 0.67$, $p = .51$), noch beim Success ($t(22) = 1.87$, $p = .07$). Nur bei der Einschätzung der Ausprägungsrate zeigt sich ein

⁶ Vier Datenwerte der Korrelationen von 1.0 mussten für die Z-Transformation ausgeschlossen werden.

signifikanter Unterschied zwischen den Gruppen ($t(22) = -3.02$, $p < .01$). Auffällig ist, dass knapp die Hälfte der Lerngruppe A die Rangreihe der Cues gemäss der Validität gut wiedergeben konnten und fast die ganze Lerngruppe B die Rangreihe der Ausprägungsrate sehr gut replizieren konnte.

Um den Unterschied zwischen dem impliziten und dem expliziten Wissen zu quantifizieren, wurde gezählt, wie häufig eine Versuchsperson den im Fragebogen angegebenen Cue auch als erster Cue im Experiment konsultiert hat. In Tabelle 19 ist der Cue aufgelistet, welche im Experiment am häufigsten (mindestens 6 mal) pro Bedingung konsultiert wurde und der Cue, welcher im Fragebogen als *bester* Cue in Bezug auf die Bedingung angegeben wurde. Falls derselbe Cue sowohl im Experiment wie auch im Fragebogen bei der jeweiligen Bedingung der Versuchsperson am nützlichsten erschien, dann kann davon ausgegangen werden, dass das implizite mit dem expliziten Wissen übereinstimmt, was als einen Übereinstimmungswert klassifiziert wird (siehe Spalte ÜW).

Tabelle 19: Erster Cue in Experiment (EXP) und bester Cue im Fragebogen (FB) in den drei Bedingungen V (Validität), D (Diskriminationsrate) und S (Success) bei den beiden Lerngruppen A (Diskriminationslernen) und Lerngruppe B (Ausprägungsrate-Lernen), wobei ÜW= Übereinstimmungswert.

Lerngruppe	V Exp	V FB	DR Exp	DR FB	S Exp	S FB	ÜW V	ÜW DR	ÜW S
A	5	5	5	5	5	5	1	1	1
A	4	4	4	4	4	4	1	1	1
A	5	5	4	4		1	1	1	0
A	5	5		4	4	5	1	0	0
A	4	4	4	3	2	4	1	0	0
A		2		2	2	2	0	0	1
A		2	3	1	3	3	0	0	1
A	5	1	5	3		1	0	0	0
A	2	5		5		5	0	0	0
A	2	5	2	5		5	0	0	0
A		5		5		5	0	0	0
A		4		2		4	0	0	0
B	2	2		2	1	5	1	0	0
B	5	4	5	5	4	5	0	1	0
B	5	4	4	4	2	4	0	1	0
B	5	2	2	2		5	0	1	0
B	5	3	2	3	3	3	0	0	1
B	5	4	2	5	2	5	0	0	0
B	1	4	3	4		3	0	0	0
B	1	3	4	3	5	4	0	0	0
B		5		5		5	0	0	0
B		5		2		1	0	0	0
B		3		3	5	3	0	0	0
B		3		1		3	0	0	0

In der Lerngruppe A gaben fünf Personen bei der Frage nach der Validität auch den Cue an, welchen sie in der Validitätsbedingung am häufigsten konsultiert haben, drei Personen zeigten eine Übereinstimmung bei dem diskriminierendsten Cue und vier Personen bei dem nützlichsten Cue gemäss Success. Bei der Lerngruppe B waren dies eindeutig weniger Personen und zwar eine Person bei der Validitätsbedingung, drei Personen bei der Diskriminationsrate und eine Person bei dem Wissen um Success.

Diskussion

Beim Testen des impliziten Wissens anhand des ersten konsultierten Cues, lässt sich einzig die Präferenz des validesten Cue 5 bei der Lerngruppe B beobachten und damit auf ein Lernen der Validität zurückschliessen. Vergleicht man die Anzahl Lerner pro Lerngruppe, dann zeigt sich auch in diesem Experiment, dass die Lerngruppe A die Validität besser lernt als die Lerngruppe B. Es scheint – gestützt durch die Daten der Nachbefragung, welche hohe Korrelationen bei der Einschätzung der Validitätsreihenfolge erbrachte – dass die Lerngruppe A die Rangreihe der Cues gemäss der Validität lernt, aber die Unterschiede zwischen den Validitäten nicht als sehr gross wahrnimmt und daher den validesten Cue nicht signifikant häufiger als die anderen Cues konsultiert.

Die Lerngruppe B zeigt zwar die Fähigkeit im Verhaltenstest die erste Information gemäss der Validität zu suchen und repliziert damit die Befunde derselben Lerngruppe im vorherigen Experiment. Aber wenn nach der Reihenfolge gefragt wird, kann diese Lerngruppe die Validität nicht explizit wiedergeben. Diese Lerngruppe B tippt bei der Diskriminationsrate häufig als erstes auf den höchst diskriminierendsten Cue, aber auch hier können die Rangreihen der Cues nicht explizit wiedergegeben werden. Es ist etwas erstaunlich, dass diese Lerngruppe die Rangreihen der Diskriminationsraten nicht besser explizit angeben kann, obwohl sie die Ausprägungsrate fast perfekt einschätzen kann und die Cues gemäss der Diskriminationsrate sucht.

Die Lerngruppe B ist eindeutig weniger gut im expliziten Einschätzen der Rangreihen der Gütemasse der Validität, Diskriminationsrate und Success. Bei der Lerngruppe A scheint das implizite und das explizite Wissen über die Cue-Güte enger zusammen zu hängen, als bei der Lerngruppe B. Newell et al. (Newell et al., 2004) fanden in ihrem experimentellen Setting (das dem paarweisen Setting der Lerngruppe A entsprach), dass 80% der Personen denselben Cue kauften, den sie als nützlichsten bewerteten. Diesen Befund finden wir in unseren Daten nicht, vor allem nicht bei den Ausprägungsrate-Lernern. Das schlechte explizite Wissen kann somit auf das Lernformat zurückgeführt werden, wobei die Alternativerklärung eines Gruppenunterschiedes nicht auszuschliessen ist, da die Gruppengrösse von zwölf Versuchspersonen doch eher an der unteren Grenze ist und sich die beiden Gruppen in ihrem Durchschnittsalter unterscheiden.

Schlussdiskussion

Im Allgemeinen lernen die Versuchspersonen die klassischen Cue-Gütemasse sehr schlecht, nur die Ausprägungsrate wird gut gelernt. Dieser Befund untermauert das Konzept des „inneren Zähler“, wie er von Hasher und Zacks postuliert wurde (1979), experimentell. Die Personen können die Ausprägungsrate explizit sehr genau wiedergeben, aber auch bestimmt diese Ausprägungsrate die Weitersuche nach einer nicht-diskriminierenden Information. Diese adaptive Suchregel, in Abhängigkeit der Art der vorliegenden Information, eröffnet Zweifel an dem Konzept der streng sequentiell geordneten Informationssuche, wie diese von Gigerenzer und Kollegen postuliert wird. Vielmehr als diese sequentielle Suchregel gesamthaft zu widerlegen, erweitert unser Konzept aber die Grundidee der „Adaptiven Toolbox“ genau um den dort wichtigen Faktor der Adaptivität. Denn ganz im Sinne dieser Adaptivität finden wir,

dass nicht nur die Wahl einer Heuristik, sondern auch die Suche selbst abhängig von der vorliegenden Information ist und die Versuchspersonen die Suche an die Umwelt anpassen. Dieser revolutionäre Befund muss jedoch in seiner Allgemeingültigkeit eingeschränkt werden. Wir finden, dass nur einige Personen tatsächlich ihre individuelle Strategie wechseln und dass die Versuchspersonen, welche anhand des klassischen Lernparadigmas die Cue-Gütemasse lernen, weder die Ausprägungsrate gut erlernen, noch diese in ihrer Informationssuche überzufällig anwenden. Es zeigt sich somit, welcher starker Einfluss das Lernformat auf das Lernen der Cue-Gütemasse Validität, Diskriminationsrate und Ausprägungsrate hat.

Die Personen, die die Gütemasse im paarweisen Setting erlernen, nehmen hauptsächlich die Validität wahr und lernen diese besser als die Personen, welche die einfachen Häufigkeiten lernen. Fraglich ist, ob es tatsächlich der Paarvergleich ist, welcher notwendig ist für diese Lernleistung oder ob es nicht vielmehr das Lernen anhand eines Feedbacks ist. Dazu müsste in einem weiteren Experiment das Ausprägungsrate-Lernen ebenfalls aktiver (z.B. durch Tippen auf die Grösse einer Stadt) und mit einem Feedback über die Richtigkeit dieser Einschätzung gestaltet werden. Enkvist, Newell, Juslin, & Olsson (2006) fanden nämlich, dass ein aktives Lernen (d.h. sie gaben das Kriterium vor und die Merkmale der Cues mussten aktiv ausgewählt werden), je nach Situation zu besserer Leistung führte und Pachur und Olsson (2008) fanden unterschiedliche Strategiewahl, je nach Art des Feedbacks (d.h. ob es absolut oder komparativ war). Solche Unterschiede in der Aktivität und im Feedback könnten auch beim Lernen und Anwenden der Validität einen Einfluss haben. Obwohl Bröder beim Erlernen der Validitäten keinen Unterschied zwischen aktivem und passivem Lernen fand (Bröder, 2000), sollte zum Zwecke der Vergleichbarkeit die Aufgabenstellung der beiden Lernformate ähnlicher gestaltet werden.

Was erstaunt ist, dass die Personen, welche bloss die Häufigkeiten anhand der Ausprägungen erlernen, die Diskriminationsrate besser lernen und anwenden können, als diejenigen Personen, welche im Paarvergleich lernen. Dies deutet darauf hin, dass tatsächlich von der Ausprägungsrate auf die Diskriminationsrate geschlussfolgert werden kann, wenn auch nur implizit. Explizit kann dieses Wissen nicht umgesetzt werden. Bei der Anwendung von Success ist mal die eine Lerngruppe besser, mal die andere, tendenziell aber zeigt die Lerngruppe A ein eher Success-orientiertes Verhalten. Dies aber hauptsächlich aus dem Grund, weil die andere Lerngruppe, die Lerngruppe B, welche die Ausprägungsrate vertiefte und auch besser lernte, im Grossen und Ganzen den Cue 2 bevorzugt. Diese unerwartete, über alle Experimente auftauchende Präferenz des Cue 2 stellt uns nach wie vor ein Rätsel. Ein möglicher Erklärungsversuch kann aus den Ergebnissen zu der Quick-Est Heuristik gezogen werden. Die Quick-Estimation Heuristik (Quick-Est), die zur Schätzung von numerischen Grössen verwendet werden soll (Hertwig, Hoffrage, & Martignon, 1999), wurde von Hausman et al. getestet (Hausmann, Läge, Pohl, & Bröder, 2007). QuickEst postuliert eine Informationssuche, die bei der Referenzklasse mit dem niedrigsten Wert startet. Hausmann et al. fanden jedoch, dass die Mehrzahl der Personen bei der mittleren Referenzgruppe mit der Informationssuche starteten. Dieser Befund könnte eine Interpretationsgrundlage dafür sein, dass der Cue 2 jeweils am häufigsten als erster Cue angeschaut wird und zwar aus dem Grund, weil der Cue 2 – vorausgesetzt man geht von vielen kleinen Städten und wenig grossen Städten aus – etwas über die mittlere Referenzgruppe aussagt.

Zusammenfassend kann gesagt werden, dass bisher angenommen wurde, dass um die ökologische Validität eines Cues zu bestimmen, es eine Referenzgruppe von Objekten benötigt, bei welchen dieser Cue diskriminiert. Newell und Kollegen kritisieren deshalb, dass diese Aufgabe, eine Hierarchie basierend auf verschiedenen Häufigkeiten von Inferenzen zu etablieren, sehr schwierig zu sein scheint (Newell et al., 2004). Wir stellen fest, dass man für die Konstruktion der Rangreihen, sowohl der Validitäten, wie auch der Diskriminationsrate diese Paarvergleiche der Referenzmengen nicht benötigt, sondern ein einfacher „innerer“ Zähler genügt. Weitere Forschung sollte aufzeigen, wie diese Ausprägungsrate mit der Validität beim Lernen interagiert und ob sie z.B. auch bei der Rekognitionsheuristik (Goldstein & Gigerenzer, 2002), im Sinne von einer Abhängigkeit der weiteren Informationssuche (je nach dem, ob zwei Städte wiedererkannt werden oder beide Städte nicht erkannt werden) eine Rolle spielt.

Literatur

- Bröder, A. (2000). Assessing the empirical validity of the "Take The Best"-heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1332-1346.
- Bröder, A. (2003). Decision making with the "adaptive toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 611-625.
- Brunswik, E. (Ed.). (1952). *The conceptual framework of psychology* (Vol. 10). Chicago: The University of Chicago Press.
- Christen, S., Hausmann, D., & Läge, D. (2006). Ist ein Gefühl für Informationsgüte lernbar? Ein Lernexperiment zu Validität, Diskriminationsrate und Success von Informationsquellen. *AKZ-Forschungsbericht* 37, Zürich: Angewandte Kognitionspsychologie.
- Enkvist, T., Newell, B., Juslin, P., & Olsson, H. (2006). On the Role of Causal Intervention in Multiple-Cue Judgment: Positive and Negative Effects on Learning. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 32(1), 163-179.
- Gigerenzer, G., & Goldstein, D. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650-669.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 506-528.
- Gigerenzer, G., Todd, P., & Gerd Gigerenzer, A. R. (Eds.). (1999). *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.
- Goldstein, D., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109, 75-90.
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 356-388.
- Hausmann, D., Christen, S., & Läge, D. (2005). Wie viel bezahlen für eine valide Information? Suchkosten als limitierender Faktor der Informationssuche. *AKZ-Forschungsbericht* 7, Zürich: Angewandte Kognitionspsychologie.
- Hausmann, D., & Läge, D. (2008). Sequential evidence accumulation in decision making: The individual desired level of confidence can explain the extent of information acquisition. *Judgement and Decision Making*, 3, 229-243.
- Hausmann, D., Läge, D., Pohl, R., & Bröder, A. (2007). Testing QuickEst: No evidence for the Quick-Estimation heuristic. *European Journal of Cognitive Psychology*, 19(3), 446-456.
- Hertwig, R., Hoffrage, U., & Martignon, L. (1999). Quick estimation: Letting the environment do the work. In P. M. T. G. Gigerenzer, & the ABC Research Group (Eds.), *Simple heuristics that make us smart*. New York, Oxford: Oxford University Press.
- Hoffrage, U., Hertwig, R., & Czienskowski, U. (2003). *The ecological rationality of Take The Best's building blocks*. Zürich: Swiss Federal Institute of Technology.
- Läge, D., Hausmann, D., & Christen, S. (2005). Wie viel bezahlen für eine valide Information? Suchkosten als limitierender Faktor der Informationssuche. *AKZ-Forschungsbericht* 7, Zürich: Angewandte Kognitionspsychologie.
- Läge, D., Hausmann, D., Christen, S., & Daub, S. (2005). Was macht einen "guten Cue" aus? Strategien der Informationssuche beim heuristischen Entscheiden unter Unsicherheit. *AKZ-Forschungsbericht* 5, Zürich: Angewandte Kognitionspsychologie.
- Martignon, L., & Hoffrage, U. (1999). Why does One-Reason Decision Making Work? A case Study in Ecological Rationality. In G. Gigerenzer, P. Todd & A. R. Group (Eds.), *Simple Heuristics that make us smart*. New York, Oxford: Oxford University Press.

- Newell, B., Rakow, T., Weston, N., & Shanks, D. (2004). Search strategies in decision making: The success of 'success'. *Journal of Behavioral Decision Making*, 17, 117-137.
- Newell, B., & Shanks, D. (2003). Take the best or look at the rest ? Factors influencing One Reason Decision Making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 53-65.
- Pachur, T., & Olsson, H. (2008). Strategiewahl in Inferenzaufgaben: Der Einfluss von absolutem und komparativem Feedback. Paper presented at the 50. Tagung experimentell arbeitender Psychologen.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 534-552.
- Rakow, T., Newell, B., Fayers, K., & Hersby, M. (2005). Evaluating Three Criteria for Establishing Cue-Search Hierarchies in Inferential Judgment. *Journal of Experimental Psychology / Learning, Memory & Cognition*, 31(5), 1088-1104.
- Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics and how can we tell? In G. Gigerenzer, P. Todd & A. R. G. Group (Eds.), *Simple Heuristics That Make Us Smart* (pp. 141-168). New York: Oxford University Press.
- Rieskamp, J., & Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica*, 127, 258-276.
- Simon, H. A. (1955). A Behavioral Model of Rational Choice. *Quarterly Journal of Economics*, 69, 99-118.

Anhang Fragebogen

1. Bitte ordne die 5 Hinweise nach ihrer Nützlichkeit zum Lösen der Entscheidungsaufgaben im Teil 2
[von *am nützlichsten* bis *am wenigsten nützlich*]
2. Bitte ordne die 5 Hinweise nach ihrer Ausprägungshäufigkeit
[von *Die in Teil 1 (Lernphase) gezeigten Städte verfügen am häufigsten über ein/eine...* bis *Die gezeigten Städte verfügen am wenigsten häufig über ein/eine...*]
3. Bitte ordne die 5 Hinweise nach ihrer Diskriminierfähigkeit
[von *Hinweis....liefert am häufigsten unterscheidbare Informationen (+ -) bzw. (- +) (von einem Städtepaar verfügt die eine Stadt über ... und die andere Stadt nicht über ...)* bis *Hinweis....liefert am seltensten unterscheidbare Informationen (+ +) bzw. (- -) (Bei den Städten verfügen entweder über ... oder beide Städte verfügen nicht über ...)*]
4. Bitte ordne die 5 Hinweise nach ihrer Trefferrate
[von *Wenn man sich auf eine unterscheidbare Information stützt, trifft dieser Hinweis am häufigsten eine korrekte Entscheidung* bis *Wenn man sich auf eine unterscheidbare Information stützt, trifft dieser Hinweis am wenigsten häufig eine korrekte Entscheidung*]

ZEITLICH DYNAMISCHE UMWELTEN

Kapitel 3: Entscheiden in zeitlich dynamischen Umwelten: Suchen und Stoppen mit unterschiedlich stabilen Cues

Einleitung

Entscheiden im Zeitverlauf

Entscheidungen brauchen Information. Diese Maxime, welche das Verhalten von Menschen in ihrer Umwelt besser macht als der Zufall, dürfte in der Forschung unbestritten sein. Und genauso unbestritten dürfte es sein, dass Menschen in der Lage sein müssen, a) sich diese Informationen zu beschaffen und b) sie so auszuwerten, dass sie auf dieser Basis die entsprechenden angemessenen Entscheidungen treffen können. Letzteres ist der traditionelle Fokus der Entscheidungspsychologie, nämlich diejenigen kognitiven Verfahren ausfindig zu machen, mit denen vorliegende Informationen in eine Entscheidung transformiert werden. In der experimentellen Forschung reicht es dazu aus, Personen mit einem (im Sinne der jeweiligen Fragestellung möglichst geschickt kombinierten) Satz von Informationen zu konfrontieren und zu beobachten, welche Entscheidungen sie treffen. Seit der Entwicklung des Information Board (z.B. in Form des *Mouselab* von Payne, 1976) und vor allem seit dem Vorschlag sequentiell arbeitender Urteilsheuristiken (z.B. Gigerenzer & Goldstein, 1996) ist außerdem auch die Frage der Informationsbeschaffung selbst immer stärker in den Fokus der Aufmerksamkeit von Entscheidungsforschern gerückt. Sequentielle Urteilsheuristiken, von denen *Take The Best* vermutlich die prominenteste ist, kombinieren eben Informationssuche und Entscheidungsfindung, indem sie Suchregel, Stoppregel und Entscheidungsregel als aufeinander bezogene Bausteine eines Gesamtverhaltens verstehen. Wie stark die Richtung einer Entscheidung von der vorausgehenden Informationssuche abhängt, konnten Läge & Hausmann (2007) zeigen, indem sie beim Einsatz eines *Information Board* für den Mehralternativenfall fanden, dass sich die Entscheidungen einer Person in derselben Situation, nur einmal mit vorgegebener und einmal mit selbst gesuchter Information, so radikal voneinander unterscheidet als würde man einen Probanden gegen einen beliebig anderen des untersuchten Samples austauschen.

Die Tradition der Forschung mit dem Information Board besteht, gleich ob mit aktiver Informationssuche oder mit vorgegebener Information, darin, der Versuchsperson eine Reihe von Durchgängen zu präsentieren, welche jeweils eine bestimmte Umwelt repräsentieren. In dieser Umwelt haben die Personen dann Entscheidungen zu treffen, und dazu können sie die angebotenen Informationen nutzen. Pro Aufgabe, Informationen und Kontextvariablen lassen sich so Erkenntnisse über den Entscheidungsprozess gewinnen. Die einzelnen Durchgänge (in der Regel einige Dutzend bis einige hundert) sind zumeist als punktuelle Messwiederholungen einer begrenzten Zahl von Variationen aufgebaut. Diese Messwiederholungen ermöglichen später eine statistische Auswertung, um die Strategien probabilistisch gegeneinander und/oder gegen den Zufall zu testen.

Diese Form der punktuellen Messwiederholung entspricht einem bewährten Standard in der experimentellen Psychologie. Was sie nicht berücksichtigen kann (und auch bei den meisten

Fragestellungen nicht zu berücksichtigen braucht), ist der Umstand, dass sich Entscheidungsverhalten nicht zu isolierten Zeitpunkten abspielt, sondern in der Realität immer in einen Zeitfluss eingebettet ist. Insofern ist die grundsätzliche Idee des *Information Board* als experimentalpsychologisches Verfahren eine Abstraktion der Wirklichkeit. Gleichwohl ist es aber auch möglich, diese Abstraktion auszuweiten und den natürlichen Zeitfluss mit einzubeziehen. Das im Folgenden zu berichtende Experiment wird das tun und eine Zeitachse einführen, auf der einzeldurchgangübergreifend Informationen gesucht und für die zu treffenden Entscheidungen verwendet werden können.

Zwar haben verschiedene Entscheidungsforscher die Zeitdimension bereits in ihre Untersuchungen mit einfließen lassen (z.B. Betsch, Haberstroh, Glöckner, Haar, & Fiedler, 2001, im Bezug auf routiniertes Verhalten; Busemeyer & Townsend, 1993, im Bezug auf Präferenzen; Kerstholt & Raaijmakers, 1997, im Bezug auf den richtigen Zeitpunkt einer Entscheidung), aber es fehlt bislang an Experimenten zum Informationssuchverhalten in Umwelten, die sich über die Zeit verändern, und zum Entscheidungsverhalten in einer Welt von sich ständig ändernden Informationen. Diese Umwelten können wir auch als „dynamisch“ bezeichnen, um sie von den „statischen“ Umwelten abzugrenzen, in denen normalerweise das Information Board mit den punktuellen Messwiederholungen Entscheidungsverhalten erhebt.

Die Einführung einer Zeitachse bedeutet für das Entscheiden auf der Basis probabilistischer Informationen zweierlei: Zum einen können Informationen aus der Vergangenheit für die aktuelle Entscheidung genutzt werden. Dazu muss man allerdings zumindest eine Ahnung haben, wie stabil denn die Zielvariablen sind (also die Alternativen, für oder gegen die man sich entscheidet oder aus denen man eine auswählt). Sind sie sehr variabel, so verliert zeitlich zurückliegende Information schnell an Validität. Neben diesem möglichen Blick in die Vergangenheit kann man aber auch in die Zukunft hinein planen und bereits jetzt nach Informationen suchen, die für eine zukünftige Entscheidung oder Wahl relevant werden können. Um das mit Erfolg zu tun, braucht man zusätzliches Wissen über die Stabilität der Prädiktorvariablen, also der probabilistischen Cues. Wenn sich ein Prädiktor über die Zeit als sehr stabil erweist, dann kann es sich lohnen, diese Information beizeiten zu besorgen und über mehrere Entscheidungen hin zu verwenden; ist er hingegen sehr variabel, so ist seine Information nur für den Moment geeignet.

Beide hiermit neu eingeführten Werte – die Stabilität der Alternativen sowie die Stabilität der Cues – hängen über die Validität der Cues in einem Punkt miteinander zusammen: Ein Cue kann nicht zugleich hochgradig valide sein, sich aber in der Stabilität stark von der der Alternativen unterscheiden. Hohe Validität setzt also ähnliche Stabilität von Prädiktor und Zielvariablen voraus. Im mittleren und unteren Validitätsbereich hingegen gilt diese Einschränkung hinsichtlich der Stabilitäten nicht mehr.

Das im Folgenden berichtete Experiment greift auf das in der Entscheidungspsychologie vertraute Börsenszenario zurück (Bröder, 2000b; Bröder & Schiffer, 2006; Läge & Hausmann, 2007; Newell, Rakow, Weston, & Shanks, 2004; Newell & Shanks, 2003; Newell, Weston, & Shanks, 2003; Rakow, Newell, Fayers, & Hersby, 2005), variiert dieses allerdings in drei Komponenten erheblich. Im Kern steht die Einführung der Zeitachse sowie des Stabilitätsfaktors als Attribut a) der Zielvariablen und b) der Prädiktorvariablen. Das Wissen einer Versuchsperson um die Stabilität der Zielvariablen ermöglicht ihr den Blick in die

Vergangenheit. Wie dabei aus unterschiedlichen Stabilitäten der Zielvariablen aus vergangenen Ausprägungen auf die derzeitige Ausprägung geschlossen wird, wurde in einem anderen Szenario bereits untersucht (zu diesem Trinkwasserszenario siehe Zurbriggen, Hausmann, Christen, & Läge, 2007). Das Wissen um die Stabilität der Prädiktorvariablen gibt der Versuchsperson zusätzlich die Möglichkeit, aus der gegenwärtigen Ausprägung dieser Prädiktorvariablen zukünftige Entscheidungen vorzubereiten.

Such- und Stoppregel in zeitlich dynamischen Umwelten

Im Rahmen der Forschung über schrittweise Urteilsheuristiken sind Such-, Stopp- und Entscheidungsregel als Bausteine einer Heuristik zu unterscheiden (so wie am Beispiel der *Take The Best*-Heuristik von Gigerenzer & Goldstein, 1996, eingeführt und in zahlreichen Folgepublikationen stärker herausgearbeitet; wird ab hier mit TTB abgekürzt). Dabei haben sich vor allem die Suchregel und die Stoppregel als für die Forschung fruchtbar erwiesen.

Bei der ***Suchregel*** geht es vor allem um die Frage, welche Qualitäten einen probabilistischen Cue attraktiv machen (und damit die Reihenfolge festlegen, mit der Cues in der Informationssuche befragt werden). Der ursprüngliche Vorschlag war die Validität (Gigerenzer & Goldstein, 1996; Gigerenzer, Hoffrage, & Kleinbölting, 1991), was verknüpft mit dem *One Reason Decision Making* zur *Take The Best*-Heuristik und bei der Integration mehrerer Cue-Informationen zu einem gewichteten Modell führt (vgl. auch Bröder, 2000a; Bröder, 2000b). In dem Moment, in dem Cues allerdings nicht immer diskriminieren, wird bei einer ressourcenaufwändigen Informationssuche auch *Success* oder *Usefulness* (also Kombinationen von Validität und Diskriminationsrate) bzw. die Diskriminationsrate selbst zu einer valablen Alternative für die Bestimmung der Qualitätsrangfolge von Cues (vgl. Läge, Hausmann, & Christen, in prep; Newell et al., 2004). Durch die Einführung einer Stabilitätsrate für die Cues kommt nun eine weitere Cue-Qualität ins Spiel: Ein besonders stabiler, gleichzeitig aber auch vernünftig valider Cue wird attraktiver, und zwar möglicherweise sogar attraktiver als ein deutlich validerer, aber instabiler Cue. Das könnte eine rein an der Validität orientierte Suchregel (so wie sie anzunehmen ist, wenn nichts als die Validität als Gütekriterium zur Verfügung steht) somit substantiell verändern. Voraussetzung dafür (und gleichzeitig Erklärungsmuster für ein entsprechendes Verhalten) ist die Intention einer Person, Cues bereits mit Blick auch auf zukünftige Entscheidungssituationen zu befragen.

Bei der ***Stoppregel*** geht es in erster Linie um die Unterscheidung zwischen *One Reason Decision Making* (stoppen nach einem diskriminierenden Cue) und *More Reason Decision Making* (fortsetzen der Suche, auch wenn bereits ein Cue diskriminierte, um eine breitere Informationsgrundlage für die Entscheidung zu besitzen). Diese beiden Formen des Entscheidens kreieren also ganz unterschiedliche Stoppregeln, und wie es scheint, haben Individuen verhältnismäßig stabile Präferenzen für das Entscheiden auf die eine oder die andere Weise (Bröder, 2000a, 2000b; Zurbriggen, Christen, Hausmann, & Läge, 2007; Zurbriggen & Läge, 2007). Allerdings finden sich starke Hinweise dafür, dass eine bestimmte (interindividuell womöglich sehr stark schwankende) Schwelle an Urteilssicherheit die Stoppregel moderiert: Sind die Validitäten der besten diskriminierenden Cues eher niedrig, tendieren die Personen zur Fortsetzung der Informationssuche, sind die Validitäten hingegen sehr hoch, dann stoppen sie gerne nach einer einzigen diskriminierenden Information (Hausmann & Läge, 2005).

Vor dem Hintergrund des *One Reason Decision Making* gewinnt die Einführung einer Zeitachse an Brisanz, denn damit liegen womöglich bereits Informationen aus vergangenen Entscheidungsdurchgängen vor und können für die aktuelle Entscheidung genutzt werden. Werden diese Informationen als in ihrer Validität noch hinreichend hoch erlebt, so müsste sich – zumindest bei Personen, die unter den entsprechenden Umständen dem *One Reason Decision Making* zugeneigt sind – eine Tendenz finden lassen, ab und zu ganz ohne neue Informationssuche zu entscheiden. Wie Hausmann, Christen & Läge (2006) in einem Experiment mit variierenden Informationskosten zeigen konnten, vermeiden es Personen normalerweise strikt, ohne jegliche Information zu entscheiden. Wenn jemand aufgrund alter Informationen eine Entscheidung trifft, ohne zuvor eine neue Suche gestartet zu haben, so werden wir dies in der Auswertung der Daten als *No New Reason Stopping Rule* (NNRStR) bezeichnen, um es der nach einem diskriminierenden Cue gestoppten Informationssuche (*One Reason Stopping Rule*, ORStR) und der auch nach einem diskriminierenden Cue noch fortgesetzten Informationssuche (*More Reason Stopping Rule*, MRStR) als dritte Stoppregel zur Seite zu stellen.

Hypothesen

Aus dem Gesagten lassen sich folgende Hypothesen ableiten. Zunächst einmal sollte die Einführung einer Zeitachse dazu führen, dass bei halbwegs stabilen Zielvariablen die Information aus der Vergangenheit für die aktuelle Entscheidung mit hinzugezogen wird. Messbare Folge ist erstens, dass die Zahl gesuchter Informationen sinkt (Hypothese 1). Zweitens sollten Personen, die sich unter den experimentellen Bedingungen der „traditionellen“ Forschung (also ohne Zeitachse) als *One Reason Decision Maker* klassifizieren lassen, eine Tendenz zeigen, zumindest in geeigneten Fällen (ausreichend hohe verbleibende Validität der Information aus der Vergangenheit) vor einer aktuellen Entscheidung ganz auf Informationssuche zu verzichten (Hypothese 2). Und drittens sollte mit dem Wissen um den Stabilitätskoeffizienten von Cues ein Cue mit besonders hoher Stabilität an Attraktivität gewinnen, d.h. häufiger gewählt werden als ohne Wissen um die Stabilität (Hypothese 3a) und – je nach Wahrnehmung der Gesamtattraktivität durch die individuelle Versuchsperson – auch häufig als erster Cue konsultiert werden als ein eigentlich deutlich validerer, aber zeitlich wenig stabiler Cue (Hypothese 3b).

Methode

Aufbau des Experimentes

Das Experiment besteht aus drei sequentiellen Bedingungen mit je 20 Durchgängen. Die erste Bedingung ist die statische Vergleichsbedingung. Hier erleben die Probanden die Ereignisse als voneinander unabhängige Durchgänge (so wie in bisherigen Entscheidungsexperimenten mit Information Board-Designs). Die Versuchspersonen können in jeder der 20 Runden („Handelstage“) in einen von zwei Optionsscheinen investieren, wobei jeweils zwei neue Optionsscheine zur Auswahl stehen. Das Information Board besteht aus einer 2x4-Matrix (zwei Alternativen x vier Cues) mit der zusätzlichen Möglichkeit, den investierbaren Betrag nicht auf eine der beiden Alternativen zu setzen, sondern als Sparbetrag direkt dem Gewinnkonto gutzuschreiben. Offen ersichtlich sind die Validitäten der Cues und die Schaltflächen für den

Informationsabruf (siehe Abbildung 10). Die Validitäten der Cues werden in ganzen Zahlen angegeben (als Treffer über 100 Durchgänge eingeführt): 90, 85, 80 und 75. Die Reihenfolge der Cues ist bei jeder Person zufällig, bleibt dann aber über die drei Bedingungen des Experiments konstant.

Den Versuchspersonen steht pro Runde ein Budget von 10'000 Franken zur Verfügung, welches sie für Investitionen und für die Informationssuche einsetzen können. Investieren sie einen Betrag in einen Optionsschein, so verdreifacht sich der Wert im Fall der positiven Entwicklung dieser Anlage, im Fall der negativen Entwicklung verfällt der Betrag.

Informationen	Trefferrate	Optionsschein X	Optionsschein Y
Indikator A	80	<input type="checkbox"/>	<input type="checkbox"/>
Indikator B	75	<input type="checkbox"/>	<input type="checkbox"/>
Indikator C	90	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Indikator D	85	<input type="checkbox"/>	<input type="checkbox"/>

Handel	
Budget	10'000 minus Suchkosten 2'000 = investierbarer Betrag 8'000
Optionsschein X	<input type="text" value="0"/>
Optionsschein Y	<input type="text" value="0"/>
Sparbüchse	<input type="text" value="8'000"/>
Auswahl bestätigen	

Übersicht	
Bisheriger Kontostand	275'000
Gewinn in den letzten 5 Tagen	89'000
(darunter 1 Tage ohne Gewinn)	
Neuer Kontostand	364'000
Anzahl vergangener Handelstage	19

Abbildung 10: Information Board der 1. Börsenbedingung

Die Probanden sind darüber informiert, dass die Wahrscheinlichkeit, dass sich ein Optionsschein positiv entwickelt, jeweils 33% beträgt und dass zwei Drittel der Optionsscheine eine negative Entwicklung haben werden. Die Versuchspersonen können jeweils in genau einen Optionsschein investieren oder den Betrag direkt auf das Gewinnkonto transferieren. In diesem Fall erhalten sie genau diesen Betrag gutgeschrieben, ohne irgendetwas zu riskieren.

Um die Entscheidung auf sicherere Füße zu stellen, stehen die vier Cues zur Verfügung. Diese können unabhängig voneinander betrachtet werden, und zwar Alternative für Alternative. (Diese Unterscheidung ist wichtig: Die Probanden betrachten nicht einen Cue mit seiner Information für alle Alternativen, sondern jeweils nur für eine Alternative. Auf diese Weise wird die Diskriminationsrate als zusätzlicher Faktor aus dem Design herausgehalten). Die Information besteht entweder aus einem roten, nach unten zeigenden Pfeil (Vorhersage, dass der Optionsschein seinen Wert verlieren wird) oder aus einem grünen, nach oben zeigenden Pfeil (Vorhersage, dass der Optionsschein an Wert gewinnen wird). Die angegebenen Validitäten zu jedem Cue sind in Bezug auf jede Bedingung korrekt, so dass sich die Probanden mit der entsprechenden Wahrscheinlichkeit auf diese Vorhersage verlassen können.

Die Kosten für das Anschauen einer Cue-Information betragen 1'000 Franken. Der Betrag für angeschaute Cues wird unmittelbar dem Rundenkonto abgezogen und kann nicht mehr investiert werden. Nach jeweils drei bis fünf Runden – das exakte Intervall wird jedes Mal neu durch Zufall bestimmt – sowie in der letzten Runde einer jeden Bedingung erhalten die Versuchspersonen Feedback über die im letzten Intervall erzielte Leistung, und zwar hinsichtlich der Höhe des im Intervall erzielten Gewinns sowie des bisher erreichten Gesamtgewinns.

Ausserdem werden sie über die Dauer des abgelaufenen Intervalls und die Anzahl Tage (im Intervall) ohne Gewinn informiert.

In der zweiten Bedingung tauchen die Personen nun in einen Zeitfluss ein (siehe Abbildung 11). Neben der Einführung der Stabilität der Alternativen bleiben die angeschauten Cue-Informationen der letzten vier Runden jeweils auf dem Information Board sichtbar. Die beiden Alternativen werden nicht mehr als Optionsscheine, sondern als Fonds bezeichnet – dies aus Plausibilitätsgründen, da nun immer das gleiche Alternativenpaar zur Auswahl steht. Den Versuchspersonen ist erklärt, dass diese beiden Fonds eine Stabilität von 80 haben, also an vier von fünf Handelstagen dieselbe Ausprägung (Wertzuwachs oder Wertverlust) besitzen wie am Vortrag.

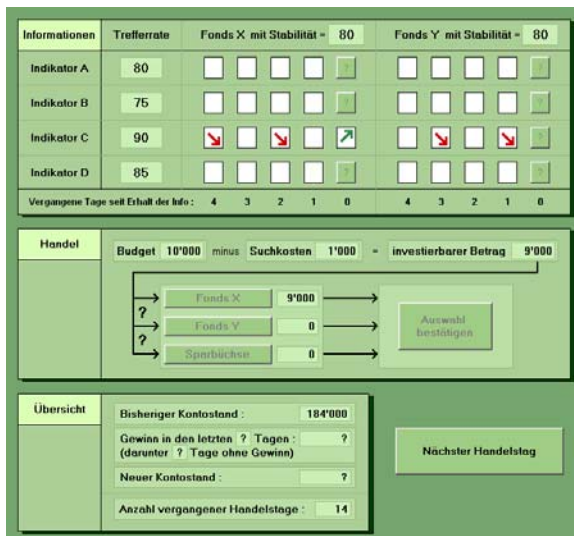


Abb. 11: Information Board der 2. Börsenbedingung

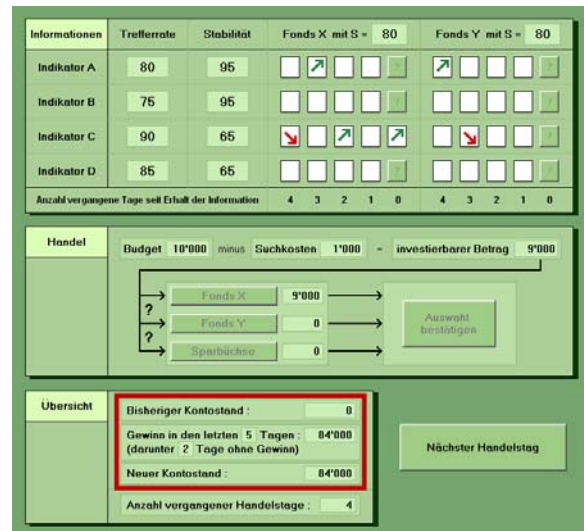


Abb. 12: Information Board der 3. Börsenbedingung

Die dritte Bedingung unterscheidet sich von der zweiten Bedingung lediglich darin, dass zusätzlich die Stabilitäten der Cues explizit auf dem Information Board angegeben werden (Abbildung 12). Die Stabilität der beiden valideren Cues ($V = .90$ bzw. $V = .85$) beträgt $.65$, die Stabilität der beiden weniger validen Cues ($V = .80$ bzw. $V = .75$) beträgt hingegen $.95$. Die Stabilitäten der Tests sind somit gegenläufig zu den Validitäten, so dass die weniger validen Cues dann attraktiver werden sollten, wenn jemand sie sich auch schon mit Blick auf künftige Handelstage anschauen möchte. Die Stabilitätsraten der Cues werden im Experiment als ganze Zahlen angegeben.

Eine detaillierte technische Beschreibung der Experimentalumgebung sowie der verwendeten Verläufe von Prädiktor- und Zielvariablen findet sich bei Christen, Zurbruggen, Hausmann & Läge (2007).

Versuchsdurchführung

Das Experiment ist in Microsoft Visual Basic 6.0 programmiert und wurde vollständig auf einem PC mit Windows XP-Betriebssystem durchgeführt. Alle Personen absolvierten die drei Versuchsbedingungen des Börsenszenarios in derselben Reihenfolge. (Da sich ja jeweils die Information über die Cue-Güte erweitert, ist es nicht sinnvoll, die Reihenfolge zu variieren. Sonst würden die Versuchspersonen womöglich ihr Wissen aus einer vorausgehenden Bedingung mit in die nächste Bedingung nehmen.)

Das Experiment wurde in Verbindung mit einem anderen Entscheidungsexperiment (in einem Trinkwasserszenario) durchgeführt. Dort waren zunächst drei Durchgänge (Versuchspersonengruppe 1) bzw. zwei Durchgänge (Versuchspersonengruppe 2) zu absolvieren, bevor die drei Bedingungen des Börsenszenarios gespielt wurden.

Stichprobe

Die Stichprobe bestand aus 60 Versuchspersonen (30 Männern und 30 Frauen), darunter 38 Psychologie-Studierende der Universität Zürich. Das Durchschnittsalter betrug 30.3 Jahre (range 19–60, $SD = 8.9$). Die Psychologie-Studierenden müssen im Rahmen ihrer Ausbildung an der Universität Zürich mindestens fünf Stunden an Experimenten teilnehmen, an welchen ist ihnen dabei frei gestellt. Die Nicht-Studierenden nahmen freiwillig teil. Unter dem Drittel mit der besten Performance wurde eine Gewinnsumme von zweimal 100 Franken verlost.

Resultate

Um den Einfluss der dynamischen Umwelt auf das Such- und Stoppverhalten zu analysieren, wurde bei den im Folgenden berichteten inferenzstatistischen Auswertungen (t-Tests für abhängige Stichproben) $\alpha = 0.05$ mit zweiseitiger Hypothesenprüfung als Grundlage für das Akzeptieren eines Mittelwertsunterschiedes gewählt.

Kaufhäufigkeiten der Cues

Kaufhäufigkeiten aller Cues

In der ersten Bedingung wurden in den 20 Runden im Mittel 40.6 Cues ($SD=14.1$) gekauft, in der zweiten Bedingung 36.9 Cues ($SD=16.7$) und in der dritten Bedingung 37.0 Cues ($SD=17.3$). Die Abnahme der Kaufhäufigkeit bei der Einführung eines zeitlichen Kontextes in Bedingung 2 gegenüber der Standardbedingung 1 ist signifikant ($t = 3.01$; $p < .01$). Dabei korrelieren die Häufigkeiten der beiden Bedingungen über die Probanden mit $r = .82$. Zwischen der zweiten und dritten Bedingung unterscheidet sich die Kaufhäufigkeiten aller Cues nicht signifikant ($t = 0.14$; $p = .89$), die Korrelation über die Probanden bleibt mit $r = .87$ hoch.

Die Versuchspersonen reagieren mit der Gesamthäufigkeit also auf die Einführung der Zeitachse (und die damit verbundene Möglichkeit, Informationen aus der Vergangenheit benutzen zu können). Die hohe Korrelation macht deutlich, dass das Gesamtmuster der Kaufhäufigkeiten über die Probanden in etwa erhalten bleibt.

Kaufhäufigkeiten einzelner Cues

Mit der zusätzlichen Einführung der Cue-Stabilitäten in Bedingung 3 ändert sich an der Gesamthäufigkeit der Cuekäufe nichts. Wie Tabelle 20 aber zeigt, reagieren die Versuchspersonen trotzdem auf die neue Situation, indem sie nämlich die beiden weniger validen, aber zeitlich sehr stabilen Cues massiv stärker berücksichtigen: Der Kauf von Cue 3 verdreifacht sich (von Bedingung 2 mit $M_2 = 3.6$ auf $M_3 = 10.9$ in Bedingung 3; $t = 5.19$; $p < .01$) und der Kauf von Cue 4 verdoppelt sich (von $M_2 = 1.8$ auf $M_3 = 3.6$; $t = -2.57$; $p < .05$). Dies geht vor allem zu Lasten von Cue 1, der in Bedingung 2 mit $M_2 = 24.8$ wesentlich häufiger gekauft

wurde als in Bedingung 3 ($M_3 = 17.1$; $t = 5.23$; $p < .01$). Der Rückgang von Cue 2 von $M_2 = 6.7$ auf $M_3 = 5.5$ erweist sich als nicht signifikant ($t = 1.43$; $p = .16$).

Tabelle 20: Summe der einzelnen Kaufhäufigkeiten in den drei Bedingungen B1, B2 und B3 (je 20 Runden); Mittelwerte und Standardabweichungen pro Versuchsperson über alle Runden.

	B1	B2	B3
Cue 1 ($V=.90$) ($S=.65$)	1718 $M_1 = 28.6$, $SD_1 = 8.3$	1489 $M_2 = 24.8$, $SD_2 = 10.3$	1024 $M_3 = 17.1$, $SD_3 = 12.1$
Cue 2 ($V=.85$) ($S=.65$)	380 $M_1 = 6.3$, $SD_1 = 7.2$	402 $M_2 = 6.7$, $SD_2 = 7.9$	330 $M_3 = 5.5$, $SD_3 = 7.7$
Cue 3 ($V=.80$) ($S=.95$)	192 $M_1 = 3.2$, $SD_1 = 4.6$	214 $M_2 = 3.6$, $SD_2 = 5.1$	653 $M_3 = 10.9$, $SD_3 = 10.4$
Cue 4 ($V=.75$) ($S=.95$)	143 $M_1 = 2.4$, $SD_1 = 4.4$	107 $M_2 = 1.8$, $SD_2 = 3.7$	214 $M_3 = 3.6$, $SD_3 = 5.7$

Bemerkenswert sind die hohen Standardabweichungen in Bedingung 3 bei Cue 1 und 3. Dies deutet auf (mindestens) zwei grundsätzlich verschiedene Muster hin, mit denen die Probanden auf die Einführung der Cue-Stabilität reagieren. Wir werden weiter unten darauf zurückkommen und diese Gruppen identifizieren.

Ein weiteres Indiz für mögliche Subgruppen sind die durchgängig niedrigeren Korrelationen der Kaufhäufigkeiten für die einzelnen Cues im Vergleich Bedingung 3 / 2 gegenüber Bedingung 2 / 1: Zwischen Bedingung 2 und 1 korrelieren die Cues folgendermaßen: Cue 1: $r = .78$, Cue 2: $r = .82$, Cue 3: $r = .61$, Cue 4: $r = .57$. Zwischen Bedingung 3 und 2 sind alle vier Werte deutlich niedriger: Cue 1: $r = .49$, Cue 2: $r = .66$, Cue 3: $r = .15$, Cue 4: $r = .41$.

Nachzutragen bleibt noch der Häufigkeitsvergleich zwischen Bedingung 1 und Bedingung 2: Hier unterscheidet sich einzig die Kaufhäufigkeit beim validesten Cue signifikant (Cue 1: $t = 4.54$; $p < .01$). Bei den weniger validen Cues hingegen gibt es keine signifikanten Unterschiede in der Kaufhäufigkeit (Cue 2: $t = 0.62$; $p = .54$; Cue 3: $t = 0.66$; $p = .52$; Cue 4: $t = 1.22$; $p = .23$). Der weiter oben berichtete signifikante Haupteffekt für die Gesamthäufigkeit zwischen Bedingung 1 und Bedingung 2 geht also voll zu Lasten des validesten Cues! In der Konsequenz muss dies bedeuten, dass die Versuchspersonen in Bedingung 2 häufiger einmal ganz ohne Cue-Kauf auskommen, und zwar primär in solchen Fällen, in denen sie in Bedingung 1 lediglich den Cue 1 angeschaut hatten. Da gleichzeitig mit dem sinkenden Mittelwert die Standardabweichung steigt (von $SD_1 = 8.3$ auf $SD_2 = 10.3$), deutet sich auch hier die Existenz (mindestens) zweier Subgruppen mit unterschiedlichem Verhaltensmuster an.

Kaufhäufigkeiten auf individueller Ebene

In Tabelle 21 werden die Cue-Präferenzen anhand der prozentualen Häufigkeiten jeder Person ersichtlich. Dabei sieht man, dass der oben beobachtete Rückgang von Cue 1 in der zweiten Bedingung vor allem auf fünf Personen zurückzuführen ist (siehe zweitletzte Spalte: Vp 30, 33, 36, 38, 43). Bei weiteren 15 Versuchspersonen kann man nicht wirklich von einer Verhaltensänderung in der Cuesuche sprechen, aber sie reduzieren dennoch die Cuekäufe von Cue 1 und wählen häufiger einmal auch andere Cues. Diese Personen zeigen im zeitlichen Kontext eine breitere Streuung der Cuesuche als in der ersten Bedingung. Zehn Personen kaufen in der zweiten Bedingung den Cue 1 häufiger als in der ersten Bedingung. Ganze 30 Personen zeigen jedoch gar keine Anpassung an die dynamische Umwelt in ihren Häufigkeiten

der Cue-Präferenzen. Die oben erwähnten Subgruppen zeigen sich somit dreigeteilt: 20 Personen, die ihre Kaufhäufigkeit von Cue 1 in der zweiten Bedingung reduzieren; 10 Personen, die diesen Cue häufiger kaufen und 30 Personen, die ihre Cue-Präferenz nicht anpassen.

Tabelle 21: Prozentuale Suche bei den einzelnen Cues (1-4) in den drei Bedingungen, sortiert nach Cue1. In den zwei letzten Spalten sind die Veränderungen von Bedingung 1 zu 2 bzw. von 2 zu 3 kategorisiert.

	1. Bedingung				2. Bedingung				3. Bedingung				Bed. 1-2	Bed. 2-3
	1	2	3	4	1	2	3	4	1	2	3	4		
4	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	NEIN	NEIN
31	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	NEIN	NEIN
32	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	NEIN	NEIN
39	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	NEIN	NEIN
41	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	NEIN	NEIN
46	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	NEIN	NEIN
47	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	NEIN	NEIN
52	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	NEIN	NEIN
53	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	NEIN	NEIN
54	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	NEIN	NEIN
58	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	NEIN	NEIN
5	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.9	0.0	0.1	0.0	NEIN	NEIN
23	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.9	0.0	0.1	0.0	NEIN	NEIN
10	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.5	0.0	0.5	0.0	NEIN	JA
25	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.5	0.0	0.5	0.0	NEIN	JA
13	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.4	0.5	0.1	0.0	NEIN	NEIN
22	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.3	0.0	0.7	0.0	NEIN	JA
8	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	NEIN	JA
33	1.0	0.0	0.0	0.0	0.4	0.4	0.2	0.1	0.4	0.4	0.2	0.1	JA	NEIN
43	1.0	0.0	0.0	0.0	0.1	0.0	0.9	0.0	0.1	0.0	0.9	0.0	JA	NEIN
6	0.9	0.0	0.1	0.0	0.9	0.0	0.1	0.0	0.7	0.1	0.1	0.1	NEIN	NEIN
4	0.9	0.1	0.0	0.0	0.8	0.2	0.0	0.0	0.7	0.1	0.2	0.0	NEIN	NEIN
12	0.9	0.1	0.0	0.0	0.8	0.2	0.0	0.0	0.6	0.0	0.4	0.0	NEIN	JA
17	0.9	0.1	0.0	0.0	0.7	0.2	0.1	0.0	0.7	0.2	0.1	0.0	NEIN	NEIN
30	0.9	0.0	0.1	0.0	0.5	0.1	0.4	0.0	0.4	0.1	0.4	0.0	JA	NEIN
56	0.8	0.1	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	NEIN	NEIN
24	0.8	0.2	0.0	0.0	0.8	0.1	0.1	0.1	0.9	0.1	0.0	0.0	NEIN	NEIN
40	0.8	0.0	0.0	0.1	0.8	0.1	0.0	0.1	0.8	0.1	0.0	0.1	NEIN	NEIN
29	0.8	0.2	0.0	0.0	0.8	0.2	0.0	0.0	0.3	0.0	0.6	0.0	NEIN	JA
18	0.7	0.2	0.0	0.1	0.9	0.0	0.0	0.0	0.5	0.0	0.5	0.0	NEIN	JA
3	0.7	0.2	0.0	0.1	0.9	0.1	0.0	0.0	0.1	0.0	0.7	0.2	NEIN	JA
2	0.7	0.0	0.3	0.0	0.8	0.0	0.2	0.0	0.7	0.0	0.2	0.1	NEIN	NEIN
9	0.7	0.2	0.1	0.0	0.7	0.3	0.0	0.0	0.8	0.1	0.2	0.0	NEIN	NEIN
1	0.7	0.3	0.0	0.0	0.7	0.3	0.0	0.0	0.7	0.3	0.0	0.0	NEIN	NEIN
21	0.7	0.3	0.0	0.1	0.7	0.2	0.0	0.0	0.7	0.2	0.0	0.0	NEIN	NEIN
44	0.7	0.3	0.1	0.0	0.6	0.3	0.1	0.0	0.6	0.3	0.1	0.0	NEIN	NEIN
51	0.7	0.2	0.1	0.1	0.5	0.2	0.2	0.1	0.5	0.2	0.2	0.1	NEIN	NEIN
19	0.7	0.3	0.1	0.0	0.5	0.3	0.2	0.0	0.4	0.4	0.2	0.1	NEIN	NEIN
36	0.7	0.0	0.3	0.0	0.3	0.2	0.2	0.2	0.3	0.2	0.2	0.2	JA	NEIN
42	0.6	0.3	0.1	0.1	0.8	0.1	0.0	0.0	0.8	0.1	0.0	0.0	NEIN	NEIN
45	0.6	0.2	0.0	0.2	0.7	0.2	0.0	0.1	0.7	0.2	0.0	0.1	NEIN	NEIN
16	0.6	0.3	0.0	0.0	0.6	0.3	0.0	0.0	0.7	0.2	0.0	0.0	NEIN	NEIN
60	0.6	0.3	0.1	0.0	0.6	0.3	0.1	0.0	0.6	0.3	0.1	0.0	NEIN	NEIN
34	0.6	0.3	0.0	0.1	0.6	0.3	0.0	0.1	0.6	0.3	0.0	0.1	NEIN	NEIN
57	0.6	0.3	0.1	0.0	0.6	0.3	0.1	0.0	0.6	0.3	0.1	0.0	NEIN	NEIN
37	0.6	0.3	0.1	0.0	0.5	0.3	0.1	0.1	0.5	0.3	0.1	0.1	NEIN	NEIN
20	0.6	0.1	0.1	0.2	0.5	0.2	0.3	0.0	0.2	0.3	0.3	0.2	NEIN	NEIN
26	0.6	0.3	0.0	0.1	0.4	0.3	0.1	0.2	0.7	0.1	0.1	0.2	NEIN	NEIN
50	0.6	0.1	0.1	0.2	0.4	0.2	0.2	0.2	0.4	0.2	0.2	0.2	NEIN	NEIN
48	0.5	0.5	0.1	0.0	0.6	0.2	0.1	0.2	0.6	0.2	0.1	0.2	NEIN	NEIN
15	0.5	0.1	0.2	0.2	0.6	0.0	0.3	0.0	0.4	0.0	0.6	0.0	NEIN	JA
35	0.5	0.4	0.1	0.0	0.5	0.4	0.1	0.0	0.5	0.4	0.1	0.0	NEIN	NEIN
59	0.5	0.3	0.2	0.1	0.5	0.4	0.1	0.0	0.5	0.4	0.1	0.0	NEIN	NEIN
7	0.5	0.3	0.1	0.1	0.4	0.3	0.1	0.2	0.5	0.2	0.2	0.1	NEIN	NEIN
28	0.4	0.3	0.1	0.1	0.4	0.3	0.1	0.1	0.4	0.5	0.1	0.1	NEIN	NEIN
49	0.3	0.2	0.2	0.3	0.5	0.1	0.1	0.3	0.5	0.1	0.1	0.3	NEIN	NEIN
55	0.3	0.0	0.7	0.0	0.4	0.0	0.6	0.0	0.4	0.0	0.6	0.0	NEIN	NEIN
27	0.2	0.0	0.4	0.4	0.1	0.1	0.6	0.1	0.2	0.2	0.3	0.3	NEIN	NEIN
38	0.2	0.2	0.2	0.3	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	JA	NEIN
11	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	NEIN	NEIN

Für eine Analyse des Einflusses der Cue-Stabilität wurden in Bedingung 2 und 3 die Häufigkeiten von Cue 1 / Cue 2 bzw. von Cue 3 / Cue 4 zusammengefasst. Dieses Vorgehen folgt der Annahme, dass die Person ihre Informationssuche nach der Validität ausrichtet, wenn sie Cue 1 oder 2 bevorzugt, und dass sie die Stabilität als Gütemaß präferiert, wenn Cue 3 und 4 in der Suche überwiegen. Um eine Verhaltensänderung zwischen der zweiten und dritten Bedingungen zu detektieren, wurden die zusammengefassten Summen der prozentualen Häufigkeiten von Cue 1 und Cue 2 bzw. Cue 3 und Cue 4 miteinander verglichen. Unterschieden sich diese um mehr als .25 in ihrer Summe, dann wurde dieses Verhalten als Änderung kategorisiert (siehe letzte Spalte in Tabelle 21). Anhand dieser Auswertung lässt sich ein veränderndes Verhalten bei neun Personen definieren (Vp 3, 8, 10, 12, 15, 18, 22, 25, 29). Diese Personen richten ihre Informationssuche

in der dritten Bedingung teilweise ausschließlich, teilweise partiell nach der Stabilität der Cues. Zusätzlich erwerben weitere 30 Personen den Cue 3 auch ab und zu. Somit zeigen sich in der dritten Bedingung noch zwei Subgruppen: 39 Personen bevorzugen den stabilen Cue und 21 Personen den validen.

Bei den 39 Personen ist jedoch unklar, ob sie die Stabilitätsrate auch richtig anwendeten oder ob die bloße Angabe der Stabilität die Attraktivität dieses Cues erhöhte. Wir nehmen an, dass die Stabilität nicht verstanden wurde, wenn der stabilste Cue nach einer Konsultation gleich

nochmals konsultiert wurde. Gegenläufig dazu gilt das Stabilitätskonzept als verstanden, wenn die Person den stabilen Cue konsultiert, ohne ihn in der Runde zuvor schon konsultiert zu haben oder ihn und gleich danach nochmals zu konsultieren. Die 39 Personen schauen den sehr stabilen Cue in der dritten Bedingung insgesamt dreimal häufiger an als in der zweiten Bedingung. Von den 653 mal, in denen er angeschaut wird, ist er 345 mal zweimal hintereinander konsultiert worden (hauptsächlich von 16 Personen) und 184 mal erfolgte dessen Konsultation ohne ihn vorher und nachher angeschaut zu haben (hauptsächlich von 14 Personen).

Suchregel

Die Häufigkeiten, mit denen die vier Cues bei den beiden Alternativen als erster, zweiter, dritter, vierter etc. angeschaut wurden, ist in Tabelle 22 dargestellt.

Tabelle 22: Häufigkeiten, mit denen die vier Cues als erstes, zweites, drittes etc. (Rangreihe 1 bis 8) angeschaut wurden; Angaben getrennt nach den drei Bedingungen bei den beiden Alternativen.

Bedingung 1	Alternative X				Alternative Y			
Rangreihe	Cue 1	Cue 2	Cue 3	Cue 4	Cue 1	Cue 2	Cue 3	Cue 4
1	885	21	34	13	177	7	16	17
2	81	111	22	19	521	46	18	18
3	7	45	24	10	31	91	19	19
4	0	18	21	7	12	24	19	16
5	0	1	4	8	4	7	6	3
6	0	3	1	1	0	2	2	6
7	0	0	0	3	0	1	3	2
8	0	1	1	1	0	2	2	0
Bedingung 2	Alternative X				Alternative Y			
Rangreihe	Cue 1	Cue 2	Cue 3	Cue 4	Cue 1	Cue 2	Cue 3	Cue 4
1	686	39	37	16	219	26	24	7
2	93	107	21	13	409	54	20	12
3	19	44	26	11	48	69	23	11
4	1	17	10	5	11	31	24	10
5	1	4	6	6	2	5	12	3
6	0	3	2	2	0	2	6	1
7	0	0	0	4	0	0	1	4
8	0	0	1	0	0	1	1	2
Bedingung 3	Alternative X				Alternative Y			
Rangreihe	Cue 1	Cue 2	Cue 3	Cue 4	Cue 1	Cue 2	Cue 3	Cue 4
1	426	50	245	59	140	20	78	22
2	100	92	53	27	238	46	142	43
3	27	31	34	12	59	35	37	15
4	5	13	8	3	23	28	29	14
5	1	2	7	4	5	8	10	1
6	0	1	1	4	0	3	4	2
7	0	0	1	2	0	0	2	4
8	0	0	0	0	0	1	2	2

Es ist zunächst auffällig, dass die Probanden eine starke Tendenz zeigen, in der Suche bei Alternative X zu beginnen. Sie steht im Information Board auf der linken Seite, so dass sich dieser Effekt durch die Gewohnheit erklärt, Dinge von links nach rechts zu lesen und zu bearbeiten. Interessant daran ist allerdings, dass der Grad der Asymmetrie von Bedingung 1 zur Bedingung 2 nachlässt: Während in Bedingung 1 das Verhältnis etwa 5:1 beträgt (885mal wird die Informationssuche mit Cue 1 von Alternative X gestartet, nur 177mal mit Cue 1 von Alternative Y), sinkt dieses Verhältnis durch die Einführung der Zeitachse auf etwa 3:1 (686mal Start mit Cue 1 der Alternative X gegenüber 219mal Start mit Cue 1 der Alternative Y). In Bedingung 3 bleibt dieses Verhältnis konstant bei 3:1 (426mal Start mit Cue 1 der Alternative X gegenüber 140mal Start mit Cue 1 der Alternative Y, und, was als substantielle Zahl hinzutritt, 245mal Start mit Cue 3 der Alternative X gegenüber 78mal Start mit Cue 3 der Alternative Y). Dass die Einführung der Zeitachse eine Verringerung der Asymmetrie nach sich zieht, macht deutlich, dass sich zumindest ein Teil der Probanden in einem Teil der Durchgänge im Start der Informationssuche an den Informationen aus den vorhergehenden Handelstagen orientiert.

Tabelle 22 zeigt zudem, dass die Einführung der Cue-Stabilität in Bedingung 3 einen sehr deutlichen Einfluss auf den zuerst angeschauten Cue hat: Während Cue 3 (Validität .80, Stabilität mit .95 aber extrem hoch) in den Bedingungen 1 und 2 keine nennenswerte Rolle spielt, wird er in Bedingung 3 nicht nur gewählt, sondern sogar oft als erster oder zweiter Cue angeschaut (immerhin etwa halb so häufig wie der valideste Cue und etwa fünfmal häufiger als in Bedingung 1 oder 2).

Es wird aber auch deutlich, dass sich bereits in Bedingung 1 nicht alle Personen in allen Fällen rational verhalten. Es gibt ja keinen Grund, mit einem anderem als dem validesten Cue die Suche zu beginnen. Trotzdem finden sich immerhin 108 Fälle (bei einer Gesamtzahl von 1200 Fällen sind das 9%), in denen die Suche mit einem weniger validen Cue begonnen wurde. Diese Fälle verteilen sich auf 13 Personen, wovon besonders die Personen 11, 27, 38, 48, 49, 50, 51 und 55 in mehr als einem Drittel aller Fälle dieses Verhalten gehäuft zeigen. Bei einer Person (Vp 49) könnte möglich sein, dass sie von vornherein intendierte, mehrere Cues aufzudecken, so dass die Frage nach dem Start für sie keine Bedeutung spielte. Die sieben anderen stoppen jedoch häufiger nach einer Information als dass sie weitersuchen. Diese Fälle deuten darauf hin, dass diese Probanden möglicherweise das Validitätskonzept nicht richtig verstanden haben.

In Bedingung 2 erhöht sich die Zahl der Durchgänge, die nicht mit dem validesten Cue begonnen wurden, nur marginal (auf 149 oder 12,4%). Bei den individuellen Häufigkeiten in Tabelle 21 zeigte sich andeutungsweise eine breite Streuung der Informationssuche. Eine starke Berücksichtigung der vergangenen Informationen (die ja zumeist den zuvor aufgedeckten validesten Cue umfassten) hätte hier nun eine höhere Zuwachsrate erwarten lassen. Für den am ehesten dazu in Frage kommenden Cue 2 (von 28 auf 65 Fälle) ist diese auch zu beobachten, allerdings auf einem insgesamt zu niedrigen Häufigkeitsniveau (der Differenzbetrag von 37 Fällen ist gerade einmal 3,1% aller Durchgänge), als dass hier von einem substantiellen Strategiewechsel in der Stichprobe gesprochen werden kann. In der dritten Bedingung wird die Informationssuche insgesamt 151mal weder mit dem validesten noch dem stabilsten Cue gestartet.

Schließlich ist zu Tabelle 22 zu bemerken, dass sich in den Häufigkeiten drei Stufen ergeben: Die eine nach dem ersten angeschauten Cue, die zweite nach dem zweiten und die dritte nach dem vierten angeschauten Cue. Vom ersten zum zweiten Cue fällt die Häufigkeit um einen Drittel, vom zweiten zum dritten Cue fällt die Häufigkeit etwa auf ein Drittel, genauso vom vierten bis zum fünften aufgedeckten Cue. Ein oder zwei Informationen scheinen also der Standardfall zu sein, in einer beträchtlichen Zahl der Fälle wird dann noch ein dritter oder vierter Cue angeschaut. Nach dem vierten ist dann aber die Informationssuche (von wenigen Einzelfällen abgesehen) beendet. Dieses Muster ist in den drei Bedingungen sehr stabil. Was man aus diesem Muster nicht herauslesen kann ist, ob es einzelne Aufgaben sind, die die Fortsetzung der Informationssuche über einen zweiten Cue hinaus nahelegen, oder ob sich Versuchspersonen in ihren Strategien interindividuell unterscheiden. Dies wird im Folgenden untersucht.

Stoppverhalten

In der Einleitung hatten wir drei verschiedene Stoppregeln definiert: Die No New Reason Stopping Rule (NNRStR) bezeichnet solche Fälle, in denen vor der jeweiligen Entscheidung keine neue Informationssuche stattfindet. Bei der One Reason Stopping Rule (ORStR) stoppt ein Proband die Informationssuche, sobald er eine diskriminierende Information gefunden hat, und bei der More Reason Stopping Rule (MRStR) setzt er die Informationssuche auch nach dem ersten diskriminierenden Cue noch fort. Da im vorliegenden Experiment die Informationen je Alternative einzeln gesucht wurden, und auch wegen der Asymmetrie der Nutzbarkeit positiver und negativer Vorhersagen, sind die möglichen resultierenden Fälle des Abbruchs einer Informationssuche vor der Auswertung der Daten diesen drei Stoppregeln zuzuordnen. Es gilt:

- Keine Information in einem Durchgang gesucht = NNRStR
- Stoppen nach dem ersten angeschauten Cue, gleich ob dieser positiv oder negativ ist = ORStR
- Fortsetzen der Informationssuche, nachdem der erste angeschaute Cue zur einen Alternative negativ war, und Stoppen nach einem angeschauten Cue der anderen Alternative (gleich ob dieser positiv oder negativ ist) = ORStR
- Suchen von mehr als einem Cue zur selben Alternative = MRStR
- Suchen nach einem Cue zur anderen Alternative, nachdem zur einen Alternative bereits ein positiver Cue gefunden wurde = MRStR

Stoppverhalten auf der Ebene der Stichprobe

Als Grundmenge für die Auswertung auf Stichproben-Ebene gilt das Stoppverhalten von 1200 Entscheidungen (20 Runden und 60 Personen) pro Bedingung. In Bedingung 1 wird insgesamt 785mal gemäß der ORStR ($M_1 = 13.1$, $SD_1 = 5.9$) gestoppt, 385mal gemäß MRStR ($M_1 = 6.4$, $SD_1 = 6.2$) und 30mal ohne eine aktuelle Information ($M_1 = 0.5$, $SD_1 = 1.6$) entschieden. In der zweiten Bedingung wird die Informationssuche 672mal gemäß ORStR beendet ($M_2 = 11.2$, $SD_2 = 5.7$), 382mal gemäß MRStR ($M_2 = 6.4$, $SD_2 = 6.4$) und 146mal ohne eine aktuelle Information ($M_2 = 2.4$, $SD_2 = 4.0$) entschieden. In der dritten Bedingung wird 622mal anhand einer Information entschieden ($M_3 = 10.4$, $SD_3 = 5.6$), 418mal anhand mehreren Informationen ($M_3 = 7.0$, $SD_3 = 6.1$) und 160mal ohne jegliche Information ($M_3 = 2.3$, $SD_3 = 4.5$).

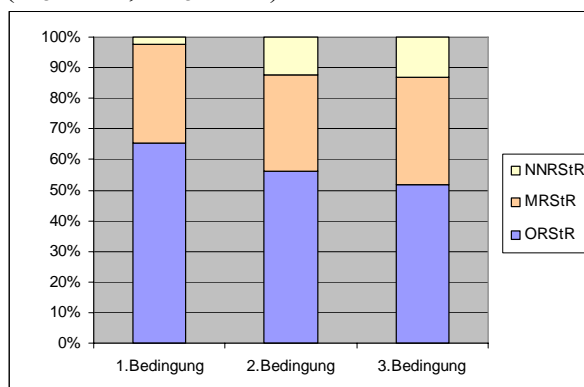


Abb. 13: Kategorisierung des Stoppverhaltens in One Reason Stopping Rule (ORStR), More Reason Stopping Rule (MRStR) und No New Reason Stopping Rule (NNRStR).

Wie Abbildung 13 zusammenfassend zeigt, wurde die ORStR in allen drei Bedingungen in mehr als der Hälfte aller Entscheidungen angewendet. Es zeigt sich jedoch eine Verschiebung von der ersten zur zweiten Bedingung: Im zeitlichen Kontext wurde weniger häufig anhand einer Information entschieden als ohne Zeitkontext ($t = 3.52$; $p < .01$), etwa gleich häufig nach mehreren Informationen gesucht ($t = 0.11$; $p = .91$), aber signifikant häufiger ($t = 4.47$; $p < .01$) ohne

jegliche Informationssuche entschieden. Die Verschiebung von der ersten zur zweiten Bedingung scheint also von ORStR zur NNRStR hin zu gehen. Es zeigen sich keine Veränderungen in den Häufigkeiten der angewendeten Stoppstrategien von der zweiten zur dritten Bedingung: ORStR mit $t = 1.65$; $p = 0.11$, MRStR mit $t = -1.34$; $p = .19$ und NNRStR mit $t = -0.78$; $p = .44$. Die anfangs auf Mittelwertsebene beobachtete Abnahme der Kaufhäufigkeit von Bedingung 1 zu Bedingung 2 entsteht also vor allem dadurch, dass im zeitlichen Kontext häufiger ohne aktuelle Information entschieden wird. Ein Blick auf die individuellen Strategien soll klären, ob die gefunden Tendenz eine allgemeine ist oder durch einige Personen verursacht wird.

Stoppverhalten auf individueller Ebene

Die Zuteilung des Stoppverhaltens einer Versuchsperson zu einer bestimmten Strategie erfolgt gemäß Binomialtest ($\alpha = 0.05$): Ab mindestens 15 Treffern bei entweder ORStR oder MRStR gilt die jeweilige Stoppregel als dominant. Bei der NNRStR nehmen wir diese Strategie als dominant an, wenn mindestens 12mal ohne aktuelle Information entschieden wurde (Da bei den 20 Runden mindestens 4mal Informationen gekauft werden mussten und der Binomialtest deswegen von einem $n = 16$ auszugehen hat).

Wie Tabelle 23 zeigt, ist in Bedingung 1 die häufigste dominante Stoppstrategie die ORStR (30 Versuchspersonen, also genau 50% der Stichprobe). Nur acht Probanden setzen signifikant am häufigsten die MRStR ein.

Tabelle 23: Anzahl Stoppverhalten, welche bei den 60 Vpn in den beiden Bedingungen gemäß One Reason Stopping Rule (O), More Reason Stopping Rule (M) und No New Reason Stopping Rule (N) klassifiziert werden. Die dominante Strategie ist fett markiert, sortiert sind die Daten nach abnehmender ORStR in der 1. Bedingung.

	1. Bedingung			2. Bedingung			3. Bedingung		
	O	M	N	O	M	N	O	M	N
53	20	0	0	20	0	0	20	0	0
14	20	0	0	20	0	0	20	0	0
13	20	0	0	20	0	0	20	0	0
32	20	0	0	20	0	0	19	1	0
10	20	0	0	19	1	0	20	0	0
52	20	0	0	19	1	0	11	9	0
58	20	0	0	19	0	1	7	13	0
8	20	0	0	17	3	0	0	6	14
39	20	0	0	16	0	4	18	0	2
46	20	0	0	14	1	5	16	1	3
33	20	0	0	12	1	7	10	1	9
55	20	0	0	10	0	10	8	0	12
41	20	0	0	4	0	16	4	0	16
23	19	1	0	20	0	0	18	2	0
22	19	1	0	19	1	0	16	4	0
25	19	0	1	16	0	4	14	1	5
31	19	0	1	14	0	6	17	0	3
36	19	0	1	13	6	1	17	3	0
17	19	1	0	12	8	0	11	9	0
3	18	0	2	13	0	7	13	0	7
6	17	3	0	17	3	0	15	5	0
54	16	4	0	17	3	0	11	9	0
56	16	4	0	15	1	4	15	2	3
40	16	4	0	12	8	0	14	4	2
24	16	3	1	10	1	9	10	1	9
5	15	0	5	16	0	4	13	1	6
27	15	5	0	16	3	1	10	10	0
4	15	5	0	9	3	8	13	2	5
43	15	0	5	6	0	14	5	0	15
30	15	4	1	4	16	0	2	17	1
18	14	6	0	19	1	0	16	4	0
12	14	4	2	5	12	3	2	14	4
37	13	7	0	9	11	0	9	11	0
48	12	8	0	13	7	0	11	9	0
51	12	8	0	10	6	4	10	9	1
47	11	9	0	13	7	0	13	7	0
50	11	9	0	13	7	0	8	10	2
29	11	9	0	10	3	7	7	3	10
38	10	10	0	17	0	3	17	2	1
60	10	10	0	10	10	0	11	9	0
2	10	10	0	8	12	0	12	8	0
11	10	0	10	5	0	15	4	0	16
45	9	10	1	11	8	1	12	1	7
15	8	12	0	7	8	5	7	9	4
42	7	13	0	12	3	5	6	12	2
16	7	13	0	8	12	0	11	9	0
9	7	13	0	7	13	0	12	7	1
44	7	13	0	6	14	0	7	13	0
21	6	14	0	10	10	0	10	10	0
1	6	14	0	7	13	0	9	11	0
34	6	14	0	6	14	0	5	15	0
19	6	14	0	2	18	0	0	20	0
20	5	15	0	5	15	0	6	14	0
49	5	15	0	3	17	0	4	16	0
57	4	16	0	4	16	0	0	20	0
59	4	16	0	2	18	0	0	20	0
26	4	16	0	1	19	0	8	12	0
28	3	17	0	6	12	2	10	10	0
7	3	17	0	2	18	0	8	12	0
35	2	18	0	2	18	0	0	20	0

Dieser Befund ist konsistent mit früheren, vergleichbaren Experimenten, jedoch ist die Asymmetrie zugunsten der ORStR wesentlich deutlicher ausgeprägt. Wenn man jedoch die hohen Validitäten berücksichtigt und annimmt, dass damit die Schwelle der Urteilssicherheit bei praktisch allen Personen überschritten sein dürfte, dann wird diese Asymmetrie zugunsten des *One Reason Decision Making* erklärlich.

Interessant ist im Kontext der vorliegenden Untersuchung aber nun, wie die Personen auf die Einführung der Zeitachse (Bedingung 2) reagieren. Insgesamt gibt es von den 60 Probanden drei, welche sich radikal auf das Minimum an notwendiger Information umstellen und eine praktisch konsequente NNRStR zeigen (mindestens viermal mussten sie ja zumindest einen Cue kaufen, um nicht zu raten). 15 weitere Personen vertrauen zumindest in einigen Fällen auf die vergangene Information (Minimum 4x NNRStR). Nur eine von ihnen (Vp 5) hatte bereits in Bedingung 1 mehrmals ein solches Verhalten gezeigt, die anderen 14 scheinen dies aufgrund der nun verfügbaren Information aus der Vergangenheit zu tun. Damit reagieren insgesamt 17 der 60 Versuchspersonen mit dem Einsetzen der NNRStR auf die dynamische Umwelt.

13 dieser 17 Personen gehören zu der Hälfte Probanden, die in der ersten Bedingung eine ORStR gezeigt hatten. Von den acht Personen, die in Bedingung 1 zur MRStR gerechnet wurden, wendet keine einzige mindestens 4x die NNRStR in Bedingung 2 an. Selbst wenn man hier das Kriterium lockert und weitere acht Personen einberechnet, die in Bedingung 1 zumindest 14- oder 13mal gemäss MRStR gestoppt hatten, findet sich nur eine einzige Person, die mehrfach die NNRStR in Bedingung 2 einsetzt (Vp 42). (Diese Person erweist sich auch noch als Wechsler hin zu einer ORStR in Bedingung 2). Die NNRStR in Bedingung 2, die mit der Zeitachse auch Informationen aus der Vergangenheit berücksichtigt, scheint also vor allem für One Reason Decision Maker eine Option zu sein. More Reason Decision Maker hingegen nutzen die Information aus der Vergangenheit nicht. Sie verwenden die Information aus der Vergangenheit auch nicht, um substantiell häufiger auf das sparsamere ORDM auszuweichen: Während die 15 Personen der „erweiterten“ MRDM-Gruppe in Bedingung 1 insgesamt 75mal die ORStR angewendet hatten, war dies in Bedingung 2 mit 71mal sogar noch etwas seltener der Fall.

Wie intraindividuell stabil diese Strategien zwischen den beiden Umwelten sind, zeigt sich daran, dass es bei den 30 One Reason Decision Makers zur Bedingung 2 hin nur einen echten „Strategiewechsler“ gibt (Vp 30 von ORStR zu MRStR), dafür aber 17, die ihrer Strategie treu bleiben. Zehn Personen adaptieren ein wenig an die dynamische Umwelt und basieren ihre Entscheidung nicht mehr ausschliesslich auf eine Information, zwei Personen betreiben konsequent NNRStR. Von den acht konsequenten More Reason Decision Makers in der ersten Bedingung, verändern sieben ihre Strategie nicht signifikant und eine Person verwendet beide in der zweiten Bedingung. Die restlichen 22 Personen verwenden in beiden Bedingungen beide Stoppstrategien uneinheitlich, wobei vor allem eine Person noch auffällt, die von einer gemischten Strategie zu einem eindeutigen ORStR in der zweiten Bedingung wechselt (Vp 38).

Über alle Auswertungen gesehen lässt sich ein klarer Zusammenhang zwischen den Subgruppen der Stoppregeln (ORDM vs. MRDM) und den Subgruppen der Informationssuche finden. Von den fünf Personen, welche den Cue 1 in der zweiten Bedingung weniger häufig kaufen,

sind vier One Reason Decision Makers in der ersten Bedingung (und einer verwendet beide Stoppstrategien gleich häufig). Von den 9 Personen, die Cue 3 in der dritten Bedingung hauptsächlich beachten, sind fünf Personen One Reason Decision Makers in Bedingung 1. (Die restlichen vier lassen sich in der ersten Bedingung keiner der Subgruppen zuordnen). Die Gruppe der More Reason Decision Makers fehlt vollständig bei denjenigen Personen, die durch die Zeitachse ihr Informationssuchverhalten maßgeblich ändern.

Diskussion

Interexperimentell stabiles Informationssuchverhalten

Die statische Vergleichsbedingung des Experiments (Bedingung 1) repliziert im Wesentlichen die Befunde, die man aus vergleichbaren Experimenten zur Informationssuche im Zwei-alternativenfall kennt: Die Personen befragen die Cues in absteigender Validitätsfolge, aber nicht immer. (In unserem Fall haben wir 9% „Ausreisser“ von dieser Suchregel). Eine Subgruppe stoppt entweder nach dem ersten positiven Cue bzw. nach den ersten Hinweisen, dass beide Optionsscheine sich negativ entwickeln werden (One Reason Stopping Rule), eine andere Subgruppe sucht trotz vorliegender Information weiter (More Reason Stopping Rule). Ferner gibt es, wie in diesen Experimenten üblich, auch diesmal eine ganze Reihe von Personen, deren Stoppverhalten bezüglich der theoretisch postulierten Strategien so uneinheitlich ist, dass man sie besser nicht klassifizieren sollte. In Such- und Stoppregel wiederholen sich also die bekannten Resultate. Dies ist nur deswegen eine Bemerkung wert, als im vorliegenden Experiment ja die Cuesuche nicht wie bei vorausgehenden Experimenten jeweils einen Hinweis für alle Alternativen gab, sondern immer nur für eine. Diese Veränderung hätte die Möglichkeit ergeben, bei den Probanden ein ganz anderes, z.B. alternativenweises Such- und Stoppverhalten zu provozieren.

Auswirkungen der Zeitachse

Am Ende der Einleitung wurden drei Hypothesen formuliert, wie sich die Einführung der Zeitachse und die Einführung der damit verbundenen Stabilitätsmaße auswirken sollten, wenn die Probanden diese neue Situation adäquat berücksichtigen.

Alle drei Hypothesen konnten durch die Resultate bestätigt werden: Der Gesamtumfang gesuchter Information sinkt in Bedingung 2 markant ab (Hypothese 1), was wir als ein Indiz deuten, dass die Probanden nun auch die Informationen aus der Vergangenheit einbeziehen und deswegen weniger neue Information benötigen. In der Tat lässt sich dieser Effekt in Bedingung 2 zu einem guten Teil auf Fälle zurückführen, in denen One Reason Decision Makers in einigen Durchgängen aufgrund der Informationen aus der Vergangenheit ein dermaßen gutes Bild von der aktuellen Situation zu haben glauben, dass sie ganz auf Informationssuche verzichten (Hypothese 2). Und mit der Einführung von Wissen über die Cue-Stabilität in Bedingung 3 verlagern die Versuchspersonen – oder zumindest ein Teil von ihnen – die Informationssuche in einer Reihe von Fällen auf die weniger validen, dafür aber stabileren Cues, und zwar sowohl in der Häufigkeit (Hypothese 3a) als auch in der Suchreihenfolge (Hypothese 3b).

Als wichtigstes Resultat zeigte sich, dass sich bei der Einführung der Zeitdimension neben der Informationssuche vor allem die Stopppregel für die Informationssuche ändert: Es werden vermehrt auf der Basis von vergangenen Informationen Entscheidungen gefällt. Es gibt zahlreiche Personen, die nicht immer eine aktuelle Information benötigen und häufiger auch mal keine Information kaufen. Jedoch wirkt sich – und das ist die wichtigste Einschränkung, die sich aus den Analysen auf individueller Ebene ergibt – die vergangene Information nur für ein knappes ein Drittel der Probanden als bedeutsam aus. Für die Mehrheit der Stichprobe ließe sich interpretieren, dass sie a) bei ihren Entscheidungen lieber auf Nummer sicher gehen und sich ungern auf vergangene (und damit unsicherere) Informationen verlassen oder b) ihre Informationssuchstrategie nicht anpassen, weil sie die Veränderungen der statistischen Umwelt nicht richtig erfassen.

Dass sich unter denjenigen Personen, die ihr Suchverhalten substantiell auf die spezifische statistische Umwelt der Zeitachse einstellen, keine einzige Person aus der Gruppe der More Reason Decision Makers in Bedingung 1 befindet, interpretieren wir durch das Modell eines erhöhten Sicherheitsbedürfnisses, welches sowohl zu MRDM als auch zu einer gewissen Ablehnung von über die Zeit doch nicht mehr so validen Cue-Informationen führt. Das Nutzen von Informationen aus der Vergangenheit oder für die Zukunft hingegen wäre dann für Personen attraktiv, welche eher an einer Minimierung des Suchaufwandes als an einer Maximierung der Entscheidungssicherheit interessiert sind.

Revival der Minimalist-Heuristik

Wir möchten an dieser Stelle gerne noch einmal auf den Befund zurückkommen, dass sich einige der One Reason Decision Makers aus Bedingung 1 in Bedingung 2 derart stark auf die Informationen aus der Vergangenheit beziehen, dass sie hinsichtlich der Cuegüte als „Minimalisten“ klassifiziert werden müssen (zumindest sich dies die beiden Vpn 41 und 43, aber man kann auch noch diejenigen sieben Personen ganz oder teilweise hinzurechnen, bei denen aufgrund zu vieler Entscheidungen ganz ohne Informationen eine Zuordnung zur ORStR nicht mehr signifikant möglich war). Diese zwei bis neun Personen sind nicht etwa wegen der Größe dieser Gruppe für eine Diskussion interessant, sondern deswegen, weil sie ja nicht nur gegen Stopppregel, sondern auch gegen die Suchregel von Take The Best verstoßen. TTB fordert, Cues in der Reihenfolge ihrer Validitäten zu befragen. Diese Reihenfolge wurde, wie wir bei den Resultaten gesehen haben, auch in Bedingung 1 in der Regel eingehalten. (Eigentlich gab es dort auch keine vernünftige Alternative dazu, da das Experiment ja die Diskriminationsrate auf 100% gesetzt hatte). In Bedingung 2 finden wir nun Personen, die systematisch dagegen verstoßen. Wenn sie dies zu ihrer Hauptstrategie machen (so wie die Probanden 41 und 43), dann ist eine Klassifikation als „Minimalisten“ zutreffend. Die Minimalist-Heuristik unterscheidet sich von TTB darin, dass die Cues zufällig und nicht nach ihrer Validität gesucht werden. Bislang lief diese Strategie in der Theorie zu Urteilsheuristiken immer unter „ferner liefen“ und wurde eher als einer der weiteren möglichen Strategien aufgelistet (z.B. Czerlinski, Gigerenzer, & Goldstein, 1999; Katsikopoulos & Martignon, 2006; Todd & Dieckmann, 2005). Jedoch gab es keine experimentellen Settings, in denen eine Anwendung dieser Strategie sinnvoll gewesen wäre. Mit der Einführung der Zeitachse liegt nun ein solches Setting vor: Wenn man bereit ist, sich ganz auf Informationen aus der Vergangenheit zu verlassen, dann verliert man an Validität, gewinnt aber an Investitionschancen (denn man gibt

ja weniger Geld für Informationssuche aus). Für risikosuchende Personen wird Minimalist also unter reinem Kosten-Nutzen-Aspekt in einer zeitlich dynamischen Welt genau so zu einer Option wie für extrem Sicherheitsbewusste, die ihr Geld lieber auf das Gewinnkonto transferieren als für Informationssuche einzusetzen (letzteres vor allem dann, wenn die Informationen aus der Vergangenheit negativ waren).

Dass Minimalist als eine Strategie in zeitlich dynamischen Welten an Bedeutung gewinnt, zeigt sich noch deutlicher in Bedingung 3 des Experiments. Gegeben einmal der Annahme, dass die Versuchspersonen die weniger validen, dafür aber sehr stabilen Cues wirklich mit Blick auf die Zukunft konsultieren, dann tun sie dies unter Preisgabe der TTB-Suchregel, welche streng absteigende Validität fordert. Ohne dass wir auf der Basis eines einzigen Experimentes bereits einen quantitativen Effekt bemühen möchten, ist die Zahl der Fälle, in denen zunächst Cue 3 oder Cue 4 konsultiert wurden, doch so erheblich, dass dies auf eine Besonderheit der präsentierten Umwelt hindeutet: Die Personen setzen nicht mehr nur auf Validität. Bislang eher ein Mauerblümchen, so könnte Minimalist also im Rahmen zeitlich dynamischer Umwelten an Bedeutung gewinnen.

Es sei angemerkt, dass diese Diskussion von Minimalist in Abgrenzung zu TTB natürlich nicht gegen TTB an sich spricht, denn diese Urteilsheuristik ist ja formuliert als optimiertes Vorgehen in Fällen, in denen Informationssuche keine Ressourcen kostet, in denen aber auf eine Begrenzung der Verarbeitungsressourcen (z.B. Arbeitsgedächtnis oder computationale Fähigkeiten) Rücksicht genommen werden muss. Wir betreten mit den zeitlich dynamischen Umwelten bei gleichzeitiger Berücksichtigung von Informationssuchkosten ein ganz anderes Feld, in dem andere Gesetzmäßigkeiten gelten könnten. Und es zeigt sich ja in den Resultaten, dass auch hier Validität immer noch die hauptsächliche Triebfeder der Informationssuche bleibt.

Schlussbemerkung

Insgesamt gesehen wäre es vermessen, den Resultaten eines ersten Experimentes in ihrem quantitativen Gefüge zueinander eine übermäßig große Bedeutung beizumessen. Mit den aufgestellten Hypothesen ist vielmehr das Feld definiert, welches bei der Berücksichtigung zeitlich dynamischer Umwelten von der Entscheidungsforschung zu bearbeiten ist. Und dass sich die Hypothesen bestätigen ließen, ist ein klares Indiz dafür, dass Menschen sensitiv auf solche zeitlich dynamische Umwelten reagieren: Zumindest einige suchen nicht nur gegenwartsorientiert Informationen, sondern berücksichtigen auch die Vergangenheit und teilweise sogar die Zukunft. Es wäre jedoch überzogen, die doch je spezifischen Restriktionen eines einzelnen experimentellen Designs unberücksichtigt zu lassen und die Resultate deswegen auf Informationssuche in zeitlich dynamischen Systemen im Allgemeinen zu erweitern. Unterstrichen haben sollte die vorgelegte Studie aber, dass hier eine interessante Forschungsaufgabe besteht, um die Entscheidungsforschung in Experiment und Theoriebildung näher an die realen Bedingungen menschlichen Entscheidens (in der Zeit nämlich) heranzuführen.

Literatur

- Betsch, T., Haberstroh, S., Glöckner, A., Haar, T., & Fiedler, K. (2001). The effects of routine strength on adaptation an information search in recurrent decision making. *Organizational Behavior and Human Decision Processes*, 84(1), 23-53.
- Bröder, A. (2000a). Assessing the empirical validity of the "Take The Best"-heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1332-1346.
- Bröder, A. (2000b). "Take the best - Ignore the rest" Wann entscheiden Menschen begrenzt rational? Lengerich: Pabst.
- Bröder, A., & Schiffer, S. (2006). Adaptive flexibility and maladaptive routines in selecting fast and frugal decision strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 904-918.
- Busmeyer, J. R., & Townsend, J. T. (1993). Decision Field Theory: A Dynamic-Cognitive Approach to Decision Making in an Uncertain Environment. *Psychological Review*, 100(3), 432-459.
- Christen, S., Zurbruggen, S., Hausmann, D., & Läge, D. (2007). Trinkwasser- und Börsenszenario: Aufbau, Ablauf und technische Spezifikation einer Experimentalumgebung zur entscheidungsvorbereitenden Informationssuche in zeitlichen Verläufen. AKZ-Forschungsbericht 43, Zürich: Angewandte Kognitionspsychologie.
- Czerlinski, J., Gigerenzer, G., & Goldstein, D. (1999). How good are simple Heuristics? In G. Gigerenzer, P. Todd & A. R. Group (Eds.), *Simple Heuristics that make us smart*. New York, Oxford: Oxford University Press.
- Gigerenzer, G., & Goldstein, D. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650-669.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 506-528.
- Hausmann, D., Christen, S., & Läge, D. (2006). Urteilssicherheit um jeden Preis? Verletzung der Stopppregel in Abhängigkeit von der Höhe der Informationssuchkosten In B. Gula, R. Alexandrowicz, S. Strauß, E. Brunner, B. Jenull-Schiefer & O. Vitouch (Eds.), *Perspektiven psychologischer Forschung in Österreich. Proceedings zur 7. Wissenschaftlichen Tagung der Österreichischen Gesellschaft für Psychologie* (pp. 45-51). Lengerich: Pabst.
- Hausmann, D., & Läge, D. (2005). How the desired level of confidence influences the stopping rule of information search. In K. Opwis & I. Penner (Eds.), *Proceedings of KogWis05. The German Cognitive Science Conference 2005* (pp. 71-76). Basel: Schwabe.
- Katsikopoulos, K. V., & Martignon, L. (2006). Naïve heuristics for paired comparisons: Some results on their relative accuracy. *Journal of Mathematical Psychology* 50(5).
- Kerstholt, J. H., & Raaijmakers, J. G. W. (1997). Decision making in dynamic task environments. In W. R. Crozier & O. Svenson (Eds.), *Decision Making: Cognitive Models and Explanations* (pp. 205-217). London: Routledge.
- Läge, D., & Hausmann, D. (2007). Der Einfluss der Informationssuche auf die Entscheidung. AKZ-Forschungsbericht 4, Zürich: Angewandte Kognitionspsychologie.
- Läge, D., Hausmann, D., & Christen, S. (in prep). Information search strategies are learnable: validity, discrimination rate, usability, and success.
- Newell, B., Rakow, T., Weston, N., & Shanks, D. (2004). Search strategies in decision making: The success of 'success'. *Journal of Behavioral Decision Making*, 17, 117-137.
- Newell, B., & Shanks, D. (2003). Take the best or look at the rest? Factors influencing one-reason decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 53-65.
- Newell, B., Weston, N., & Shanks, D. (2003). Empirical tests of a fast and frugal heuristic : not everyone "takes-the-best". *Organizational Behavior and Human Decision Processes*, 91, 82-96.
- Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, 16, 366-387.
- Rakow, T., Newell, B., Fayers, K., & Hersby, M. (2005). Evaluating Three Criteria for Establishing Cue-Search Hierarchies in Inferential Judgment. *Journal of Experimental Psychology / Learning, Memory & Cognition*, 31(5), 1088-1104.
- Todd, P. M., & Dieckmann, A. (2005). Heuristics for ordering cue search in decision making. In L. K. Saul, Y. Weiss & L. Bottou (Eds.), *Advances in neural information processing systems* (Vol. 17). Cambridge, MA [et al.]: MIT Press.
- Zurbruggen, S., Christen, S., Hausmann, D., & Läge, D. (2008). Adaptivität der Informationssuche in statischen und in dynamischen Umwelten. AKZ-Forschungsbericht 46, Zürich: Angewandte Kognitionspsychologie.
- Zurbruggen, S., Hausmann, D., Christen, S., & Läge, D. (2007). Die zeitliche Dynamik von Umwelten: Heuristisches Wahlverhalten am Beispiel eines Trinkwasser-Szenarios mit und ohne Zeitachse. AKZ-Forschungsbericht 41, Zürich: Angewandte Kognitionspsychologie.
- Zurbruggen, S., & Läge, D. (2007). Ressource Zeit statt Geld: Ein Vergleich zweier Szenarien zur Entscheidungsforschung. AKZ-Forschungsbericht 45, Zürich: Angewandte Kognitionspsychologie.

Kapitel 4: Die zeitliche Dynamik von Umwelten: Heuristisches Wahlverhalten am Beispiel eines Trinkwasser-Szenarios mit und ohne Zeitachse

Einleitung

Das Leben spielt sich in einem Zeitverlauf ab, und so auch das Entscheiden. Nicht nur, dass es gute und weniger gute Zeitpunkte für eine Entscheidung gibt (vgl. Kerstholt & Raaijmakers, 1997), sondern auch die Informationen, welche einer Entscheidung zugrunde liegen, ändern sich über die Zeit. Deswegen sollte die Suche nach den probabilistischen Hinweisen auf die beste wählbare Alternative auch möglichst eng an den Zeitpunkt der Entscheidung selbst gekoppelt sein, wenn man eine maximale Validität von Prädiktorvariablen (Cues) für eine Zielvariable (Entscheidungsalternative) erreichen will.

Die bisherige experimentelle Entscheidungsforschung hat dies, soweit sie sich überhaupt mit Informationssuche als einem aktiven Prozess zur Vorbereitung einer Entscheidung beschäftigt hat, dadurch realisiert, dass Informationssuche und Entscheidung auf denselben Zeitpunkt gelegt wurden (Bröder, 2000a, 2000b, 2002, 2003, 2005; Bröder & Schiffer, 2003, 2006; Gigerenzer, Todd, & ABC Research Group, 1999; Hausmann & Läge, 2005; Läge, Hausmann, & Christen, 2005; Läge, Hausmann, Christen, & Daub, 2005; Newell, Rakow, Weston, & Shanks, 2004; Newell & Shanks, 2003; Newell, Weston, & Shanks, 2003; Payne, 1976; Payne & Bettman, 2000; Payne, Bettman, & Johnson, 1988; Payne, Bettman, & Johnson, 1993; Rieskamp & Hoffrage, 1999). Die dabei eingesetzten Information Boards waren dann entsprechend so strukturiert, dass es sich wiederholende Durchgänge gab, in denen jeweils Informationen gesucht und auf der Basis dieser Informationen entschieden werden konnte. Der nächste Durchgang hatte dann (von der Information her gesehen) nichts mehr mit dem vorausgehenden Durchgang zu tun.

Dies ist natürlich eine Abstraktion der Wirklichkeit, eine sinnvolle freilich, wenn man – wie hier im Kontext der Erforschung schrittweiser Entscheidungsheuristiken – etwas über die Reihenfolge der Informationssuche, über das Beenden dieser Suche beim Vorliegen bestimmter Informationen sowie über die Entscheidungsregel bei den dann vorliegenden Informationen wissen möchte. Die drei Schritte gelten als die Bausteine einfacher Entscheidungsheuristiken (vgl. z.B. (Gigerenzer et al., 1999) und erhielten deswegen eine gewisse Aufmerksamkeit in der bisherigen experimentellen Forschung.

Auf der Basis dieser Ansätze lässt sich nun die experimentelle Wirklichkeit ein wenig näher an die Lebenswirklichkeit heranrücken, wenn man berücksichtigt, dass der Zeitpunkt der Informationssuche nicht immer der Zeitpunkt der Entscheidung ist. Häufig verfügen wir ja bereits aus der Vergangenheit über Informationen, welche wir in eine aktuelle Entscheidung mit einfließen lassen können. Diese Informationen sind allerdings, sofern sie probabilistischen Charakter besitzen, nicht mehr ganz so valide wie eine aktuell gesuchte Information zu einer oder mehreren Entscheidungsalternativen (vorausgesetzt einmal, Prädiktorvariable und Zielvariable sind nicht hundertprozentig über die Zeit stabil). Also werden wir – sofern wir mit der eingeschätzten verbleibenden Validität der Informationen aus der Vergangenheit nicht mehr zufrieden sind – unser Wissen durch eine aktuelle Informationssuche ergänzen. Dies ist ein ganz normales Verhalten in den zeitlich dynamischen Umwelten unseres Alltags und lässt

sich durchaus in der experimentellen Entscheidungsforschung modellieren. Man muss nur eine Zeitachse in die bestehenden Information Boards einbauen. Hinzutretende Untersuchungsvariablen wären dann beispielsweise der Aktualitätsgrad der Information oder – für die Abschätzung der Situation in einer zeitlich dynamischen Umwelt enorm wichtig – der Stabilitätsgrad einer Alternative.

Ausserdem wird in einer Umwelt mit zeitlicher Verlaufsperspektive auch die Zeit selbst zu einer wichtigen Ressource. Auch wenn wir uns gerne Zeit als infiniten Strahl vorstellen, so ist sie für Organismen doch grundsätzlich begrenzt. Auch für die Erledigung bestimmter Aufgaben gibt es immer spezifische zeitliche Begrenzungen. Hat man mehrere Dinge zu tun, so können diese miteinander um die durch die grundsätzliche Begrenzung entstandene Ressource Zeit konkurrieren. Das ist bereits im Entscheidungsprozess der Fall: Zeit, die jemand für die Informationssuche aufwendet, kann er nicht mehr für einen anderen Bereich einsetzen. Deswegen wird er bestrebt sein, den Prozess der Informationssuche möglichst effizient zu gestalten.

Bislang wurde in der Entscheidungsforschung dieser grundsätzlich bestehende Aufwand bei der Informationssuche durch monetäre Kosten experimentell implementiert. Das ist sinnvoll, um Aufwand und Erreichtes (der Wert von Alternativen, für die sich eine Versuchsperson entschieden hat) möglichst direkt miteinander in Beziehung setzen zu können. Grundsätzlich ist aber auch denkbar, dass man ein Experiment gestaltet, in dem die Zeit selbst als Ressource für die Informationssuche herangezogen wird. Eine solche Idee ist beispielsweise in einem Experiment realisiert, in dem ein Information Board möglichst schnell zu bearbeiten ist (z.B. Rieskamp & Hoffrage, 1999, 2006; Kerstholt, 1994). Aber natürlich braucht man hier nicht unbedingt ein „Echtzeitszenario“ einzuführen.

Im Folgenden werden wir ein Experiment vorstellen, welches eine Zeitachse implementiert und dabei auch bei der Ressource ganz auf die Zeit setzt. Dieses Experiment wird als zweite experimentelle Neuerung einen Stabilitätskoeffizienten für die möglichen Entscheidungsalternativen einführen und so den Versuchspersonen die Möglichkeit geben, die verbleibende Validität zeitlich zurückliegender Informationen korrekt einzuschätzen. Das Experiment versteht sich als Erweiterung der bisherigen Untersuchungen zur Erforschung schrittweiser Urteilsheuristiken. Aus diesem Grund wollen wir – bevor wir das Experimentaldesign selbst vorstellen – noch kurz überlegen, welche Auswirkungen die Einführung einer Zeitachse auf die drei wichtigsten Bausteine dieser Heuristiken (Suchregel, Stopppregel, Entscheidungsregel) haben.

Suchregel

Der wichtigste Index für die Qualität eines probabilistischen Cues ist seine Validität. Nicht zuletzt deswegen schlagen beispielsweise auch Gigerenzer & Goldstein (1996) vor, in einer einfachen Entscheidungsheuristik die Cues nach absteigender Validität zu befragen, bis eine diskriminierende Information gefunden wurde. Sofern man über Cues verfügt, welche unmittelbar eine Information über alle Entscheidungsalternativen geben, können diese Cues in einem grösseren Set von Alternativen nicht immer zwischen allen Alternativen diskriminieren. Vor allem wenn die Suche nach Information mit Aufwand verbunden ist, kann es ratsam sein, neben der Validität auch die Diskriminationsrate in die Überlegungen nach der optimalen

Reihenfolge der zu suchenden Information einzubeziehen. Dies ist beispielsweise im Success-Koeffizienten (Newell et al., 2004) und im Usefulness-Koeffizienten (Hoffrage, Hertwig, & Czienskowski, 2003) der Fall. In beiden führt eine Verrechnung von Validität und Diskriminationsrate zum besten Index für die Festlegung einer Suchregel bei Informationskosten. Diese Koeffizienten sind aber nur in Fällen relevant, in denen die Cues nicht immer diskriminieren. Ansonsten gilt weiterhin die Validität als beste Suchregel.

Im Kontext der Einführung einer Zeitachse wird diese Validität nun massgeblich durch die Stabilität der Alternativen beeinflusst, zumindest was das Heranziehen von Cue-Informationen aus der Vergangenheit anbelangt: Bei einer über die Zeit perfekt stabilen kann man sich genauso gut auf einen gleich validen Indikator aus der Vergangenheit verlassen wie auf einen aus der Gegenwart. Je geringer die Stabilität jedoch wird, umso geringer wird auch die Vorhersagekraft vergangener probabilistischer Cue-Information für die aktuelle Ausprägung des Cues. Und je mehr Zeit (hier Anzahl Möglichkeiten für eine Zielvariable, ihre Ausprägung zu wechseln) verstrichen ist, umso mehr sinkt die Validität eines Cues gegen Basisrate einer zufälligen Vorhersage.

Dieser Zusammenhang lässt sich experimentell implementieren, indem man Alternativen unterschiedlicher Stabilität in ein Setting einbaut und indem es einen Zeitverlauf gibt, also eine variierende Anzahl von Möglichkeiten, dass eine Alternative ihre Ausprägung ändert. Versuchspersonen sollten dann eher bei einer stabileren Alternative auf vergangene Information setzen als bei einer instabilen Alternative.

Beenden der Informationssuche

Bei der Stopppregel geht es in erster Linie um die Unterscheidung zwischen *One Reason Decision Making* (ökonomisches Stoppen nach dem ersten Cue) und *More Reason Decision Making* (Fortsetzen der Suche, auch wenn bereits eine Cue-Information vorliegt, um eine breitere Informationsgrundlage für die Entscheidung zu gewinnen). Wie es scheint, haben Individuen verhältnismässig stabile Präferenzen für das Entscheiden auf die eine oder die andere Weise (Bröder, 2000a, 2000b; Zurbriggen, Christen, Hausmann, & Läge, 2007a, 2007b; Zurbriggen & Läge, 2007).

Vor dem Hintergrund des *One Reason Decision Making* gewinnt die Einführung einer Zeitachse an Brisanz, denn damit liegen womöglich bereits Informationen aus vergangenen Entscheidungsdurchgängen vor und können für die aktuelle Entscheidung genutzt werden. Demnach müsste sich – zumindest bei Personen, die unter den entsprechenden Umständen dem *One Reason Decision Making* zugeneigt sind – eine Tendenz finden lassen, ab und zu ganz ohne neue Informationssuche zu entscheiden (was in einem Setting mit unabhängigen Einzeldurchgängen strikt vermieden wird, vgl. Hausmann, Christen & Läge, 2006). Wenn jemand aufgrund alter Informationen eine Entscheidung trifft, ohne zuvor eine neue Suche gestartet zu haben, so werden wir dies in der Auswertung der Daten als „No New Reason Stopping Rule“ (NNRStR) bezeichnen, um es der nach einem Cue gestoppten Informationssuche („One Reason Stopping Rule“, ORStR) und der auch nach einer Cue-Information noch fortgesetzten Informationssuche („More Reason Stopping Rule“, MRStR) als dritte Stopppregel zur Seite zu stellen.

Entscheidungsverhalten

Im *One Reason Decision Making* ist das Fällen einer Entscheidung eigentlich eine Trivialität: Man hat eine einzige diskriminierende Cue-Information auf dem Tisch, und natürlich entscheidet man sich in diese Richtung. Im *More Reason Decision Making* benötigt man eine Integrationsregel, nach der die unterschiedlichen Informationen zu gewichten sind. Beide Strategien sind sowohl in statischen wie in zeitlich dynamischen Umwelten in gleicher Weise anzuwenden.

Allerdings kommt, sofern man etwas über die Stabilität der Alternativen weiss, ein weiterer Faktor hinzu. Aus der Ausprägung der Alternative in der Vergangenheit lässt sich bei hinreichender Stabilität ja auch etwas über die gegenwärtige Ausprägung sagen. Dieser Faktor wirkt allerdings nur, sofern man die tatsächliche Ausprägung der Alternative zum vorausgehenden Zeitpunkt auch kennt. Dies kann in einem Experimentalsetting durch geeignetes Vorhalten des Feedbacks natürlich ausgeschlossen werden, sofern es im Sinne der Beantwortung der Hypothesen nicht sinnvoll ist, dass die Probanden über dieses Wissen verfügen. Doch auch wenn Personen nur eine wage Ahnung über die Ausprägung einer zeitlich stabilen Alternative haben, sollte dies ihr Vertrauen in die richtige Entscheidung bei dieser Alternative erhöhen, verglichen zu einer wenig stabilen Alternative, über die dieselbe aktuelle Information vorliegt.

Mit der Einführung einer Zeitachse ergibt sich zudem die Möglichkeit, nicht nur eine Entscheidung auf einen späteren Zeitpunkt zu verschieben, sondern auch die Informationssuche. Beides wird man vor allem dann machen, wenn die Rahmenbedingungen, überhaupt eine günstige Entscheidung treffen zu können, aus den Informationen der jüngsten Vergangenheit heraus als ungünstig beurteilt werden und wenn man sich gleichzeitig Hoffnung auf Besserung der Rahmenbedingungen in der Zukunft machen darf. Voraussetzung für eine solche Strategie ist, dass das Aufschieben keine substantiellen Ressourcen kostet.

Methode

Aufbau der Experiments

Um den Unterschied zwischen einer statischen, aus Einzelentscheidungen bestehenden Umwelt und einer zeitlich dynamischen Umwelt miteinander verbundener Entscheidungssituationen experimentell zu implementieren, wurde als Information Board ein Trinkwasserszenario geschaffen: Die Versuchsperson erhält die Aufgabe, für die fiktive mexikanische Firma AquaMex Trinkwasser abzufüllen, welches jeweils zeitgleich mit zwei Tanklastwagen von zwei Quellen der Umgebung herangefahren wird. Da dieses Wasser sich nur zum Teil als Trinkwasser eignet (bei einem Drittel der Ladungen), steht neben der Abfüllstation auch eine Teststation mit vier durchführbaren Tests (den probabilistischen Hinweiscues) zur Verfügung. Diese vier Tests haben unterschiedliche Validitäten (0.90, 0.85, 0.80, 0.75). Sie können einzeln pro Wasserladung durchgeführt werden, und zwar beliebig viele in beliebiger Häufigkeit. Es ist dann zu entscheiden, ob man a) genau eine der beiden Wasserlieferungen zu Trinkwasser abfüllen möchte (die Abfüllanlage ist in ihrer Kapazität begrenzt) oder b) beide Wasserlieferungen wegschüttet (d.h. den Bauern in der Umgebung zur Bewässerung ihrer Felder zur Verfügung stellt), weil man in keine der Lieferungen genügend Vertrauen hinsicht-

lich der Wasserqualität hat. Entscheidet man sich für Variante a), so wird die entsprechende Lieferung in Trinkwasserflaschen abgefüllt. Eine staatliche Behörde überprüft am Ende des Prozesses, ob das Wasser tatsächlich als Trinkwasser in den Handel kommen darf. Sie gibt gelegentlich (im Experiment alles drei bis fünf Durchgänge mit Abfüllung) Auskunft darüber, wie viele der Trinkwasserabfüllungen qualitativ in Ordnung waren. Die Versuchsperson soll im Experiment möglichst viele Wasserlieferungen abfüllen, die den Test der staatlichen Behörde bestehen.

Für dieses Szenario steht der Versuchsperson ein Zeitbudget zur Verfügung, welches in der Summe 240 Stunden beträgt. Auf diese Weise werden Kosten in das Experiment eingeführt: Das Abfüllen einer Wasserlieferung in der Abfüllstation kostet sie 10 Stunden Zeit, das Durchführen jedes einzelnen Tests 1 Stunde. Das Wegkippen einer Lieferung übernimmt der Fahrer des Tanklastzugs, so dass die Versuchsperson hierfür keine Zeit aufzubringen hat. Die Zeit verstreicht mit jeder Aktion, und eine Versuchsbedingung ist beendet, wenn alle 240 Stunden aufgebraucht sind.

In diesem Szenario lassen sich nun drei Versuchsbedingungen implementieren: eine mit unabhängigen Einzelereignissen und zwei mit zeitlich verbundenen Ereignissen, die letztere davon mit zusätzlichem Wissen über die Stabilität der Zielvariablen.

Die Versuchsbedingung der „zeitlosen“ Umwelt dient als Ausgangspunkt des Experiments und gleichzeitig als Vergleich für die Versuchsbedingungen mit einer Zeitachse (im weiteren Text wird diese Bedingung als T1 für Trinkwasser-Szenario 1 bezeichnet). Hier stammen die Wasserladungen jeweils aus zwei von vielen möglichen Quellen aus der Umgebung (siehe Abbildung 14). Die jeweilige exakte Herkunft ist nicht bekannt, so dass keine Schlussfolgerungen aus früher durchgeführten Tests möglich sind. Die Personen werden mit folgendem Text in dieses Szenario T1 eingeführt:

Das Quellwasser wird jeweils von verschiedenen Wasserquellen geholt und mit Tanklastwagen zu AquaMex geliefert, wo es eine Wasseranalysestation und eine Abfüllstation gibt. [...] Falls du die Wasserqualität einer Tankladung für gut befindest, kannst du mit dieser Tankladung zur Abfüllstation fahren. [...] Falls du der Meinung bist, dass keine der beiden Tankladungen trinkbares Quellwasser enthält, kannst du beide Lastwagen zu den benachbarten Landwirten schicken. Dies bringt dir mit Sicherheit keine Einnahmen, dafür verbrauchst du jedoch keine Abfüllzeit.

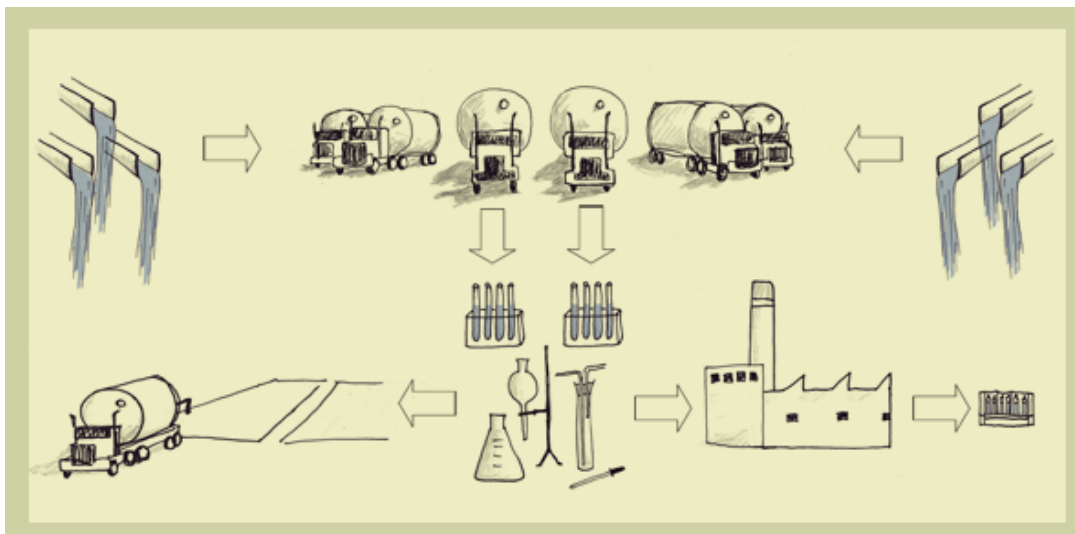


Abbildung 14: Illustration zur Einführung in die Bedingung des Trinkwasser-Szenarios mit Einzelereignissen (T1)

In der zweiten Bedingung (Trinkwasser-Szenario 2, kurz T2) stammen die Tankladungen in jeder Runde aus denselben beiden Wasserquellen X und Y (siehe Abbildung 15). Dies impliziert die gewünschte zeitliche Abhängigkeit der Zielvariablen. Die durchgeführten Tests der letzten vier Zeitpunkte bleiben auf dem Information Board sichtbar, so dass die Versuchspersonen bei der möglichen Verwendung von Testinformationen aus der Vergangenheit nicht an einer Begrenzung des Arbeitsgedächtnisses scheitern. Die Probanden wissen über die beiden Quellen lediglich, dass sich deren Wasserqualität von Zeit zu Zeit ändern kann und dass beide im Mittel in einem Drittel der Fälle brauchbares Trinkwasser liefern. Alle anderen Umstände bleiben gleich wie in Bedingung T1. Folgender Auszug aus dem Instruktionstext für T2 zeigt den Unterschied zu T1 auf:

In deiner Arbeit musst du dich jedoch ein wenig umstellen, denn es stehen für die Anlieferung nur noch zwei Wasserquellen zur Verfügung. Da die Wasserquellen immer dieselben zwei sind und sich die Wasserqualität nicht willkürlich ändert, bleiben die Informationen der Tests der letzten vier Zeitpunkte nutzbar. Die ankommenden Tanklastwagen bilden zwei Schlangen, die eine von Wasserquelle X kommend, die andere von Wasserquelle Y kommend. Du musst dich jeweils entscheiden, ob du eine der beiden Ladungen abfüllen willst und falls ja, welche.

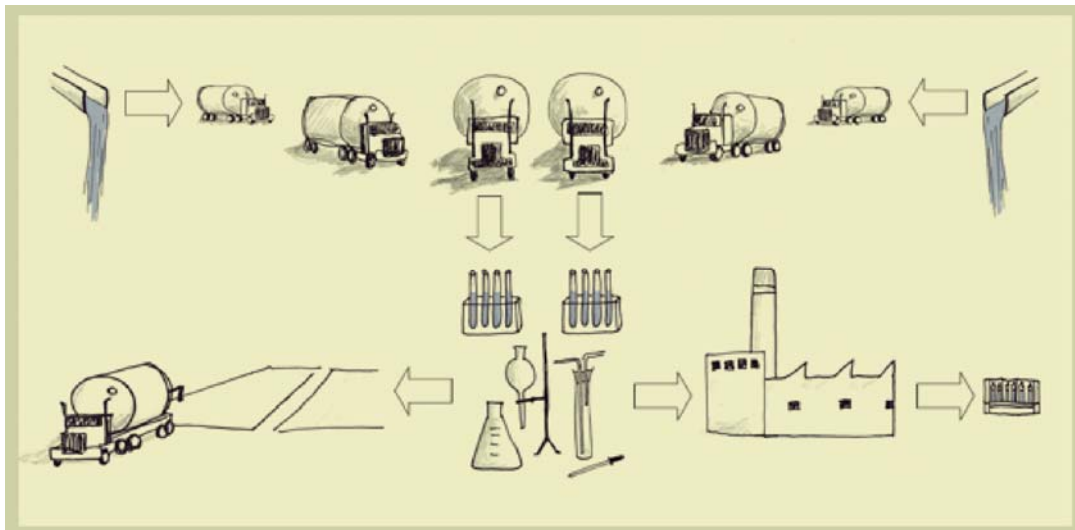


Abbildung 15: Illustration zur Einführung in die Bedingung des Trinkwasser-Szenarios mit zeitlich dynamischen Ereignissen (T2)

In der dritten Bedingung (Trinkwasser-Szenario 3, kurz T3) wird nun die Stabilitätsrate der Wasserqualität aus den beiden Quellen X und Y explizit mitgeteilt, ansonsten entspricht die Bedingung T2. Die Stabilitätsrate der Wasserqualität ist bei den beiden Wasserquellen jedoch unterschiedlich: Quelle X hat eine Stabilität von .60 und die Quelle Y von .90.

In diesen drei Bedingungen erweitert sich jeweils das Wissen der Versuchsperson um einen Faktor. Deswegen ist es sinnvoll, die drei Bedingungen in genau dieser Reihenfolge den Versuchspersonen vorzulegen. Eine detaillierte Beschreibung aller technischen Parameter des Experiments und eine Herleitung der hinter den Prädiktor- und Zielvariablen stehenden Datenverläufe findet sich bei Christen, Zurbriggen, Hausmann & Läge (2007). Alle den Probanden mitgeteilten Validitäten, Stabilitätsraten und Basisraten sind in jeder Versuchsbedingung korrekt umgesetzt, so dass sich die Versuchspersonen auf diese Angaben verlassen können.

Operationalisierte Hypothesen

Hinsichtlich der Stoppregel und hinsichtlich der Entscheidungsregel erwarten wir Unterschiede im Verhalten, wenn sich die Umwelt von einer statischen zu einer zeitlich dynamischen wandelt. Allerdings enthält ein erstes Experiment mit Einführung einer entsprechenden Zeitachse natürlich einen gewissen Explorationsfaktor. Es kann z.B. nicht vorhergesagt werden kann, ob Personen allein schon durch die Möglichkeit, Informationen aus der Vergangenheit nutzen zu können, darauf reagieren werden oder doch erst mit einem Wissen über die Stabilität der Zielvariablen. Es wird also jeweils festzustellen sein, ob sich bereits ein Unterschied von T1 zu T2 ergibt oder erst von T2 zu T3.

Die erste wichtige in dieser Hinsicht zu überprüfende Variable ist die Anzahl durchgeführter Tests pro Durchgang. Mit der Möglichkeit, Informationen aus der Vergangenheit heranzuziehen, sollte sich die Zahl der aktuell durchgeführten Tests verringern (Hypothese 1).

Des weiteren eröffnet die Zeitachse die strategische Möglichkeit des Abwartens: Wenn die Tests zu beiden Quellen in jüngster Vergangenheit eine schlechte Wasserqualität angezeigt haben, dann gibt es ja die Möglichkeit, sogar ganz ohne Test beide Lieferungen den Bauern zu überlassen (Hypothese 2 = Anzahl der Durchgänge, in denen beide Lieferungen weggeschüttet werden, steigt gegenüber T1). Das ist der Extremfall der Variante. Im moderaten Fall würde vielleicht noch ein kontrollierender Test durchgeführt, bevor man die Lieferungen wegkippt. In der Summe sollte dies darin resultieren, dass durch die „gesparte“ Zeit die Anzahl der mit den 240 Stunden spielbaren Durchgänge durch die Einführung der Zeitachse zunimmt (Hypothese 3).

Das bedeutet eine Veränderung der Stoppregel: Entscheiden ohne Informationssuche (nicht zu erwarten in statischen Szenarien) sollte ab T2, spätestens jedoch bei T3 eine valable Alternative werden (Hypothese 4). Ausserdem sollte sich die Attraktivität der zusätzlichen Suche bestätigender Information über einen ersten diskriminierenden Hinweis hinaus verringern; die More Reason Stopping Rule sollte also seltener gewählt werden (Hypothese 5).

Diese Veränderungen erwarten wir allerdings nicht für die gesamte Stichprobe, sondern primär für Subgruppen, die sich anhand ihres Stoppverhaltens in T1 voneinander unterscheiden lassen: Wer in T1 mehrheitlich *One Reason Decision Making* betrieb, der sollte nun die Tendenz zeigen, immer einmal wieder ganz ohne neue Informationen zu entscheiden (Hypothese 6). Personen, die in T1 zum *More Reason Decision Making* neigten, sollten hingegen die Cue-Informationen aus der Vergangenheit mit einbeziehen und deswegen in der jeweils aktuellen Informationssuche auch häufiger einmal mit einer One Reason Stopping Rule auskommen (Hypothese 7).

Hinter den Zielvariablen (und damit auch hinter den Prädiktorvariablen) stehen bestimmte Zeitverläufe, die den in T3 offengelegten Stabilitätsraten entsprechen. In T1 sind diese irrelevant, weil die Durchgänge logisch unverbunden sind. Mit dem Wissen um die zeitliche Verbindung in T2, spätestens aber mit dem Wissen um die Stabilitätsraten in T3, und unterstützt durch die Informationen aus der Vergangenheit, haben die Versuchspersonen die Möglichkeit, derartige Verläufe zu entdecken. Dies sollte darin resultieren, dass die Wahrscheinlichkeit dafür steigt, dass dieselbe Alternative, wenn sie beim vorausgehenden Durchgang zum Abfüllen gewählt wurde, unmittelbar wieder gewählt wird (Hypothese 8).

Mit der Einführung der Stabilitätsfaktoren in T3 wird die Quelle Y gegenüber der Quelle X attraktiver. Es ist deswegen zu erwarten, dass sich die Informationssuche in T3 im Vergleich zu T2 stärker auf Quelle Y verlagert (Hypothese 9). Dasselbe erwarten wir für die Anzahl der Entscheidungen (Hypothese 10), wobei dieser Effekt aber nicht gross sein kann, da ja beide Alternativen gleich häufig (ein Drittel der Fälle) eine gute Wasserqualität bieten. Eine Möglichkeit zu einem grösseren Effekt bietet hingegen die Betrachtung je zwei unmittelbar aufeinanderfolgender Durchgänge: Wenn die stabilere Alternative Y im vorausgehenden Durchgang als positiv bewertet wurde (also zur Abfüllung kam), dann sollte sie im aktuellen Durchgang in T3 eine höhere Wahrscheinlichkeit der Testdurchführung (Hypothese 11) und vor allem auch eine höhere Wahrscheinlichkeit der wiederholten Abfüllung (Hypothese 12) haben als noch in T2. Für die Quelle X (um deren weitaus niedrigere Stabilität die Versuchsperson ja in T3 weiss) erwarten wir das nicht.

Mit der Einführung der Stabilitätsinformation wird für die Probanden in T3 auch das Risiko kalkulierbar, wenn sie die Durchführung neuer Tests reduzieren und eher auf vergangene Information setzen. Das könnte sich im Suchverhalten und im Stoppverhalten ausdrücken, und zwar in einer Asymmetrie zwischen der stabileren Alternative (hier ist der Test vom vorausgehenden Durchgang noch mehr wert) und der weniger stabilen Alternative: Da die Versuchspersonen ja in der Regel den validesten Test durchführen dürften und dieses Resultat damit vom letzten Durchgang vorliegt, müsste es also gehäuft für die stabile Alternative Y zu Fällen kommen, in denen dieser valideste Test überhaupt nicht konsultiert wird (Hypothese 13). Personen sollten sich häufiger einmal ganz ohne aktuellen Test für die stabilere Alternative entscheiden als für die weniger stabile Alternative (Hypothese 14). Und Personen, die in T1 mehrheitliches MRDM betreiben, könnten vornehmlich bei der stabileren Alternative Testresultate aus vorherigen Durchgängen mit einbeziehen und somit mit nur einem einzigen zusätzlichen aktuellen Test genügend Sicherheit für ihre Entscheidung gewinnen (Hypothese 15).

Als Konstante über alle drei Bedingungen sollte sich die Suchregel ergeben: Hier erwarten wir keine Abweichungen von der Validitätsfolge (Hypothese 16). Auch die Leistung der Versuchspersonen (gezählt in korrekt abgefüllten Lieferungen) sollte nicht durch den Wechsel in den Umwelten an sich steigen (denn die Zahl brauchbarer Lieferungen ist ja durch das Design gleichmässig begrenzt). Ein Zusammenhang zwischen Leistung und Bedingung könnte lediglich über eine grössere Anzahl gespielter Durchgänge moderiert sein (Hypothese 17), denn mit mehr Durchgängen erhöht sich ja die Chance des Findens zusätzlicher guter Wasserlieferungen.

Diese 17 Hypothesen sind nicht alle völlig unabhängig voneinander, sondern stellen in einigen Fällen Facetten desselben Grundphänomens dar. Fast jede von ihnen kann jedoch eintreten, ohne dass eine andere Hypothese dafür erfüllt sein müsste. Ein Hauptaugenmerk wird für die Hypothesen 1 – 8 darauf zu legen sein, ob, wenn sie eintreten, dies bereits bei T2 (nur Wissen über die Abhängigkeit) oder erst bei T3 (zusätzliches Wissen über die Stabilität) der Fall ist.

Versuchsdurchführung und Versuchspersonen

Insgesamt absolvierten 30 Versuchspersonen (Alter $M = 29.4$, $SD = 8.0$; 14 Männer und 16 Frauen) die drei Bedingungen des Trinkwasserszenarios. Aus Gründen einer einheitlichen, aufeinander aufbauenden Instruktion wurden die drei Bedingungen jeweils in derselben Reihenfolge absolviert: zuerst die „zeitlose“ Bedingung (T1), dann die Bedingung, in welcher die Umwelt des Experiments durch die Zeitdimension erweitert wird (T2) und zuletzt die Bedingung mit Stabilitätsangaben zu den Alternativen (T3). Unter dem besten Drittel der Versuchspersonen wurde ein Gewinn von CHF 100.- verlost, um die Probanden zu einer guten Gesamtleistung zu motivieren, ohne sie aber zu extrem riskanten Grenzstrategien zu verleiten.

Resultate

Bei den im Folgenden berichteten inferenzstatistischen Auswertungen (t-Tests für abhängige Stichproben oder, wo zwischen Personengruppen getestet wurde, auch für unabhängige Stichproben) wurde $\alpha = 0.05$ mit zweiseitiger Hypothesenprüfung als Grundlage für das Akzeptieren eines Mittelwertsunterschiedes gewählt.

Leistung

Die Anzahl „guter“ Trinkwasserabfüllungen als zentrales Mass für die Leistung einer Versuchsperson im Experiment steigt von Bedingung zu Bedingung. In T1 werden im Mittel 11.8 Punkte ($SD = 2.0$, Range = 9 – 15), in T2 12.6 Punkte ($SD = 2.8$, Range = 9 – 18) und in T3 12.9 Punkte ($SD = 2.1$, Range = 9 – 16) erzielt. Die Leistungssteigerung von der ersten Bedingung (T1) zur zweiten Bedingung (T2) ist knapp signifikant ($t = 2.2$; $p < .05$), jedoch nicht signifikant ist der marginale Anstieg der Leistung von Bedingung 2 zu Bedingung 3 ($t = 0.6$; $p = .57$). Von T1 zu T2 verschlechtern sich 7 Personen in ihrer Leistung, 6 Personen stagnieren und 17 Personen verbessern ihre Leistung. Von T2 zu T3 verschlechtern sich 12 Personen in ihrer Leistung, 4 Personen halten die Leistung und 14 verbessern ihre Leistung. Die erbrachte Leistung korreliert in jeder Bedingung signifikant ($p < .01$) mit der Anzahl absolvierter Runden (in T1 $r = .69$, in T2 $r = .75$ und in T3 $r = .50$). Das heisst, dass Personen, welche eher mit Entscheidungen zuwarten, in der Tendenz bessere Entscheidungen treffen.

Suchregel

Verteilung der Informationssuche auf die einzelnen Cues

Die Häufigkeiten, mit welcher die Cues konsultiert werden, korrespondieren in allen drei Bedingungen mit der Reihenfolge der Validität der vier Cues (Tabelle 24), und zwar gleich ob man sich die Gesamtzahl pro Bedingung oder den Durchschnitt pro Durchgang anschaut (letzterer wird mit ausgewertet, da sich die Zahl der Durchgänge über die Personen und über die Bedingungen stark voneinander unterscheiden konnte). Die Suchregel entspricht also (wie nicht anders zu erwarten) einer lexikografischen Suchstrategie, und zwar mit einem ganz klaren Sprung zwischen dem validesten und dem zweitvalidesten Cue.

Tabelle 24: Anzahl gesuchter Cues pro Versuchsperson in den drei Bedingungen T1 („zeitlos“), T2 (mit Zeitkontext) und T3 (mit Stabilitätsangaben zu den Alternativen) über alle Runden und pro Runde.

alle Runden	T1	T2	T3
Cue 1 (V=.90)	M = 38.9, SD = 7.5	M = 42.3, SD = 15.0	M = 37.8, SD = 13.1
Cue 2 (V=.85)	M = 16.5, SD = 13.3	M = 17.2, SD = 16.8	M = 12.7, SD = 13.6
Cue 3 (V=.80)	M = 8.0, SD = 6.9	M = 6.2, SD = 7.5	M = 4.7, SD = 6.5
Cue 4 (V=.75)	M = 5.8, SD = 7.7	M = 4.2, SD = 7.1	M = 3.1, SD = 5.4
Summe	M = 69.2, SD = 21.6	M = 69.9, SD = 32.0	M = 58.3, SD = 28.4
pro Runde	T1	T2	T3
Cue 1 (V=.90)	M = 1.5, SD = 0.3	M = 1.4, SD = 0.5	M = 1.3, SD = 0.5
Cue 2 (V=.85)	M = 0.7, SD = 0.6	M = 0.6, SD = 0.6	M = 0.5, SD = 0.5
Cue 3 (V=.80)	M = 0.3, SD = 0.3	M = 0.2, SD = 0.3	M = 0.2, SD = 0.3
Cue 4 (V=.75)	M = 0.3, SD = 0.4	M = 0.2, SD = 0.3	M = 0.1, SD = 0.2
Summe	M = 2.8, SD = 1.2	M = 2.3, SD = 1.2	M = 2.0, SD = 1.1

Bemerkenswert sind sicherlich die ab Cue 2 stark ansteigenden Standardabweichungen. Dies weist darauf hin, dass sich zwei Untergruppen von Versuchspersonen bilden, die einen, die nach dem validesten Cue die Suche abbrechen, und die anderen, die danach noch regelmässig weitersuchen. Auch die stark ansteigende Standardabweichung von Cue 1 von Bedingung 1 zu Bedingung 2 ist bemerkenswert. Offenbar weitet sich mit dem Wechsel zu einer Umwelt mit Zeitachse die Verhaltensvarianz der Versuchspersonen hinsichtlich des Durchführens des ersten Tests enorm aus. Wir werden zu einem späteren Zeitpunkt der Auswertung feststellen, dass dies damit zu tun hat, dass eine Reihe von Personen in einer substantiellen Zahl von Fällen ganz ohne Informationssuche entscheidet (also weniger Cue 1 konsultiert) und dass sich eine andere Gruppe noch stärker als zuvor auf Cue 1 konzentriert.

Ein extrem seltenes Ereignis ist das Ignorieren des validesten Cues in einem Durchgang: Wenn Information gesucht wird, dann ist er auch praktisch immer mit dabei. Mit Blick auf die aufgestellte Hypothese 13 ist für diese Fälle interessant, wie häufig sich die Probanden in T3 mit Tests, aber ohne Cue 1 für eine der beiden Alternativen entschieden haben. Insgesamt handelt es sich dabei um 40 Fälle, verteilt auf acht Versuchspersonen: 25mal fiel dabei die Entscheidung auf die stabilere Alternative Y (M = 3.1, SD = 3.8), 15mal auf die weniger stabile Alternative X (M = 1.9, SD = 1.9). Im t-Test wird dieser Unterschied nicht signifikant ($t = 1.1$, $p = .29$), wobei ein signifikanter Unterschied im Blick auf die nur acht Probanden, die sich überhaupt in Umgehung von Cue 1 für das Abfüllen von Wasser entscheiden, statistisch ohnehin problematisch wäre.

Anzahl durchgeführter Tests

Insgesamt werden in der „zeitlosen“ Bedingung T1 in etwa gleich viele Informationen wie in der zweiten Bedingung mit Zeitdimension T2 gesucht ($M_{T1} = 69.2$; $M_{T2} = 69.9$; $t = 0.2$; $p = .83$). Kommen bei der dynamischen Umwelt jedoch noch zusätzlich Stabilitätsinformationen hinzu (T3), werden signifikant weniger Informationen gesucht ($M_{T3} = 58.3$; $t = 5.0$; $p < .01$). Diese Abnahme ermöglicht den Versuchspersonen, im Mittel einmal mehr Wasser abzufüllen (Zeitbedarf 10 Stunden) als in T1 und T2. Pro Runde betrachtet – und das ist im Sinne der Hypothesen der wichtigere Wert – sinkt die Zahl der durchgeführten Tests von $M_{T1} = 2.8$ über $M_{T2} = 2.3$ auf $M_{T3} = 2.0$. Die Unterschiede zwischen je zwei aufeinander folgenden Bedingungen sind signifikant ($t_{T1-T2} = 4.1$, $p < .01$; $t_{T2-T3} = 4.0$, $p < .01$).

In allen drei Bedingungen wird häufiger bei der Alternative X als bei der Alternative Y nach Information gesucht. In T1 werden 1234 Tests zur Alternative X durchgeführt und nur 841 zur Alternative Y ($t = 5.9$; $p < .01$). In T2 wird 1113mal bei der Alternative X und 981mal bei der Alternative Y ein Test konsultiert ($t = 3.7$; $p < .01$). In T3 allerdings verschwindet der Unterschied fast, mit 890 Tests zu Alternative X und 860 Tests zu Alternative Y ($t = 0.4$; $p = .69$).

Tabelle 25 zeigt die Wahrscheinlichkeit, mit der zu einer Alternative ein Test durchgeführt wird, wenn in der unmittelbar vorhergehenden Runde diese Tankladung für das Abfüllen des Wassers gewählt wurde. Diese Wahrscheinlichkeit ist für Alternative X (0.78, 0.79, 0.71) jeweils deutlich höher als für Alternative Y (0.31, 0.41, 0.62). Allerdings verschwindet diese Asymmetrie in T3 zum grössten Teil, indem der Wert für Quelle Y von 0.41 (in T2) auf 0.62 (T3) drastisch ansteigt.

Tabelle 25: Häufigkeit der Suche bei derjenigen Alternative, welche in der unmittelbar vorhergehenden Runde (Zeitpunkt t-1) gewählt wurde. Ausgewertet über alle Runden.

	T 1 Alternative X	T 1 Alternative Y	T 2 Alternative X	T 2 Alternative Y	T 3 Alternative X	T 3 Alternative Y
Total Wahl	342	162	256	237	258	276
Total Suche, wenn Alternative zu (t-1) gewählt wurde	268	51	201	96	184	171
Anteil	0.78	0.31	0.79	0.41	0.71	0.62

Beenden der Informationssuche

Wie wir bei der Tabelle 24 bereits gesehen haben, werden mit zunehmenden Stabilitätsangaben weniger Cues pro Runde gesucht. Zusätzlich zeigt sich, dass auch mehr Runden absolviert werden, wenn man etwas über die Stabilität der Umwelt weiss: In T1 werden im Durchschnitt 26.6 Runden ($SD = 5.0$) absolviert, in T2 sind es 31.2 Runden ($SD = 5.3$) und in T3 31.5 Runden ($SD = 9.5$). Die Anzahl Runden unterscheidet sich zwischen den ersten beiden Bedingungen signifikant, zwischen den zeitlich dynamischen Bedingungen hingegen nicht mehr ($t_{T1-T2} = 8.80$; $p < .01$; $t_{T2-T3} = 0.19$, $p = .86$).

Bei der Einführung der Hypothesen haben wir in Anlehnung an die bereits bestehende Forschung eine Klassifizierung des Stoppverhaltens in drei Kategorien vorgeschlagen: No-New-Reason Stopping Rule (Entscheiden ohne Hinzuziehen von neuen Informationen), One-Reason Stopping Rule und More-Reason Stopping Rule. Da im vorliegenden Experiment die Informationen je Alternative einzeln gesucht wurden, sind die möglichen resultierenden Fälle des Abbruchs einer Informationssuche vor der Auswertung der Daten diesen drei Stopppregeln zuzuordnen. Es gilt:

- Keine Information in einem Durchgang gesucht = NNRStR
- Stoppen nach dem ersten angeschauten Cue, gleich ob dieser positiv oder negativ ist und Fortsetzen der Informationssuche mit der anderen, nachdem der erste angeschauten Cue zur einen Alternative negativ war, und Stoppen nach einem angeschauten Cue der anderen Alternative (gleich ob dieser positiv oder negativ ist) = ORStR
- Suchen von mehr als einem Cue zur selben Alternative und Suchen nach einem Cue zur anderen Alternative, nachdem zur einen Alternative bereits ein positiver Cue gefunden wurde = MRStR

Auf der Mittelwertsebene zeigt Tabelle 26 (rechte Spalte *Sum*) eine Tendenz zu häufigerem Anwenden von ORStR in allen drei Bedingungen. In der ersten „zeitlosen“ Bedingung T1 stoppen die Probanden insgesamt etwas häufiger gemäss einer ORStR ($M = 13.8$, $SD = 12.4$), als gemäss einer MRStR ($M = 12.4$, $SD = 9.3$), jedoch nicht signifikant ($t = 0.4$; $p = .73$). Zudem entscheiden sich die Versuchspersonen in T1 wie angenommen kaum je ohne eine Information ($M = 0.3$, $SD = .61$; sieben der zehn Fälle finden in der allerersten oder allerletzten Runde statt). In T2 wird ebenfalls häufiger anhand einer einzigen Information entschieden ($M = 14.5$, $SD = 11.5$) als anhand mehrerer Informationen ($M = 12.6$, $SD = 10.8$), aber auch hier nicht signifikant ($t = 0.5$; $p = .64$). Dafür entscheiden sich die Probanden in T2 in vielen Runden ohne aktuelle Informationssuche ($M = 4.0$, $SD = 7.9$). In T3 wird wiederum häufiger anhand einer einzigen Information entschieden ($M = 15.8$, $SD = 11.2$) als anhand mehrerer ($M = 10.4$, $SD = 9.9$), aber auch das nicht signifikant ($t = 1.5$; $p = .15$). Häufig wird auch hier ohne jegliche Information entschieden ($M = 5.2$, $SD = 9.8$).

Die Anzahl der Fälle, in welchen ohne eine neue Information entschieden wird, unterscheiden sich in T1 und T2 signifikant ($t = 2.5$; $p < .05$). Jedoch wird in beiden Bedingungen gleich häufig gemäss ORStR ($t = 0.4$; $p = .66$) gestoppt und auch gleich häufig gemäss MRStR ($t = 0.3$; $p = .74$). Die Häufigkeiten der Stoppregeln in den beiden dynamischen Bedingungen T2 und T3 unterscheiden sich weder in der Anzahl Fälle, in welchen keine Information gesucht wird ($t = 1.1$; $p = .28$), noch in Bezug auf die ORStR ($t = 1.0$; $p = .31$). Jedoch sinkt die Häufigkeit des MRStR von T2 nach T3 hin signifikant ($t = 2.3$; $p < .05$).

Das Stoppverhalten auf individueller Ebene steht in Tabelle 26 im Mittelpunkt. Dabei wurde die jeweils dominante Stoppstrategie einer Person gemäss Majoritätsregel festgelegt (fett eingefärbt), ist also nicht per Binomialtest abgesichert und dient deswegen der deskriptiven Darstellung und nicht der Auszählung klar inferenzstatistisch identifizierbarer Strategien. In T1 werden danach 14 Personen als ORStR-Anwender klassifiziert und 16 Personen stoppen mehrheitlich gemäss einer MRStR. 19 der 30 Probanden behalten ihre Majoritäts-Stoppregel über das gesamte Experiment hin bei (10 aus der MRStR-Gruppe und 9 aus der ORStR-Gruppe). Vier der 10 aus der MRStR-Gruppe stoppen jedoch in T2 und T3 markant häufiger auch einmal nach einem einzigen Cue, während eine umgekehrte Tendenz bei den 10 konstanten der ORStR-Gruppe angehörenden Personen nicht zu erkennen ist. Im Gegenteil, sie entscheiden häufiger einmal ganz ohne Informationen (zwei Personen in T2 und fünf in T3), was sich aus der MRStR-Gruppe lediglich Vp15 in T3 ein paar Mal erlaubt.

Bei den Majoritätswechslern setzt sich dieser Trend fort, und zwar intraindividuell in entsprechend verstärkter Form (Strategiewechsel eben): Sechs ursprünglich die MRStR bevorzugende Personen wechseln entweder gleich in T2 oder spätestens in T3 zur ORStR, während vier Personen, die in T3 die ORStR bevorzugten, mit der Einführung der Zeitachse zur NNRStR als Majoritätsfall wechseln (Vp14 allerdings in T2 erst halbherzig). Die Tendenz ist also klar: Personen, die in einer Umwelt ohne Zeitachse die MRStR bevorzugen, zeigen bei der Möglichkeit, auch Informationen aus der Vergangenheit heranzuziehen, mehrheitlich eine Tendenz, mit einer neu gesuchten Information zufrieden zu sein. Zur NNRStR hingegen wechselt niemand aus dieser Gruppe. Bei vielen Personen, welche in einer Umwelt ohne Zeitachse zum *One Reason Decision Making* tendierten, zeigt sich eine klare Tendenz, nun häufiger einmal ganz auf die Informationssuche zu verzichten und lediglich auf die Information aus

der Vergangenheit zu vertrauen. Keine Regel ohne Ausnahme allerdings: Vp21, die mit *One Reason Decision Making* startet, verwendet in T3 schliesslich doch die MRStR.

Tabelle 26: Anzahl Runden, in denen eine Versuchsperson a) ohne neue Information eine Entscheidung trifft (NNRStR) b) sich mit einer Information entschied (ORStR) oder c) bei bereits vorliegender Information noch weiter sucht (MRStR). In den Spalten sind die 30 Versuchspersonen angegeben, sortiert nach absteigender Differenz zwischen ORStR und MRStR in Bedingung T1. Die Majorität des Verhaltens legt jeweils die Strategie dieser Person (fett gedruckt) fest.

	14	29	5	11	23	22	10	25	3	8	4	21	27	24	17	12	6	13	16	18	28	26	2	15	1	7	20	9	19	30	Sum
T1 NNRStR	2	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	1	1	0	0	0	10
T1 ORStR	45	29	28	28	28	27	26	26	26	24	21	18	15	15	12	7	7	7	9	5	0	4	1	1	3	0	0	1	0	1	414
T1 MRStR	0	0	0	0	0	0	2	2	3	3	7	9	11	12	17	15	16	18	20	18	16	23	21	21	24	21	21	23	24	26	373
T2 NNRStR	24	22	2	22	0	1	0	5	0	0	24	0	6	11	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	120
T2 ORStR	26	13	29	7	29	29	34	28	31	28	4	28	18	18	19	9	2	23	12	8	0	11	12	1	0	0	9	6	1	0	435
T2 MRStR	1	0	0	0	1	0	2	2	1	3	7	3	8	7	14	11	25	11	21	17	22	21	19	19	30	29	21	28	28	29	380
T3 NNRStR	42	20	0	17	4	0	0	24	15	0	12	0	0	12	0	1	0	0	0	0	3	0	1	5	1	0	1	0	0	0	158
T3 ORStR	23	13	29	11	23	29	25	29	35	44	16	9	15	16	20	14	14	29	19	19	0	6	9	1	5	1	5	9	0	6	474
T3 MRStR	0	0	0	0	0	0	5	0	1	0	1	20	10	0	10	8	11	0	14	11	22	22	22	16	23	25	22	20	27	22	312

In den 158 Durchgängen, in denen in T3 ohne eine Information entschieden wurde, wurde 81mal eine der beiden Wasserlieferungen abgefüllt. Dabei wurde 35mal die Alternative X und 46mal die Alternative Y gewählt. Vergleicht man diese Werte bei den zwölf Probanden, die für mindestens eine der beiden Alternative einen Score >0 haben, so ergibt sich kein signifikanter Unterschied bei den beiden Alternativen ($t = 0.7$, $p = .49$).

Für die 16 Probanden, die in T1 mehrheitlich MRDM gezeigt hatten, sind die Anzahl Fälle bei den beiden Alternativen in T3 zu vergleichen, in denen anhand einer einzigen neuen Information für eine Alternative entschieden wurde. Dies ist bei der Alternative X 135mal der Fall und bei der Alternative Y 165mal. Dieser Unterschied ist nicht signifikant ($t = 0.2$, $p = .86$).

Entscheidungsverhalten

Wie Tabelle 27 zeigt, wird in T1 die Alternative X deutlich häufiger gewählt als die Alternative Y ($t = 7.2$; $p < .01$). Diese extreme Asymmetrie verschwindet in T2 weitgehend ($t = 1.2$; $p = .24$) und kippt in T3 sogar (jedoch nicht signifikant; $t = 0.4$; $p = .72$).

Tabelle 27: Häufigkeit der Wahl der drei möglichen Entscheidungsalternativen: Summen über alle Probanden, individuelle Mittelwerte und Standardabweichungen für alle drei Experimentalbedingungen

	T1	T2	T3
Alternative X	342 M = 11.4, SD = 2.7	256 M = 8.8, SD = 2.4	258 M = 8.6, SD = 4.6
Alternative Y	162 M = 5.4, SD = 2.3	237 M = 7.9, SD = 2.9	276 M = 9.2, SD = 4.7
Felder bewässern	293 M = 9.7, SD = 4.1	433 M = 14.4, SD = 5.2	410 M = 13.7, SD = 9.0

Der entscheidende Wechsel im Entscheidungsverhalten findet zwischen T1 und T2 statt: Alle drei Wahlhäufigkeiten, die der Alternative X die der Alternative Y und die des Felderbewässerns, unterscheiden sich zwischen T1 und T2 signifikant voneinander: Alternative X wird weniger gewählt ($t_{T1-T2} = 5.3$, $p < .01$), Alternative Y dafür häufiger ($t_{T1-T2} = 5.0$, $p < .01$), und ebenfalls häufiger entscheiden sich die Personen dafür, beide Wasserladungen wegzuschütten

($t_{T1-T2} = 7.0$, $p < .01$). Zwischen T2 und T3 hingegen gibt es keinerlei signifikanten Veränderungen ($t_{T2-T3} = 0.3$; $p = .79$ bei Alternative X, $t_{T2-T3} = 1.4$; $p = .18$ bei Alternative Y und $t_{T2-T3} = 0.6$; $p = .57$ beim Wegschütten beider Ladungen).

Tabelle 28 zeigt die Häufigkeit der unmittelbaren Wiederholung der Wahl einer Alternative in zwei aufeinander folgenden Durchgängen. Diese ist zunächst in T1 für Alternative X höher als für Alternative Y (was schlicht an der grösseren Basishäufigkeit der totalen Wahl liegen dürfte). Dieser Effekt reduziert sich in T2, da der Unterschied in der Basishäufigkeit verschwindet. Das eigentlich zu bemerkende Resultat ist jedoch der markante Anstieg beider Wahrscheinlichkeiten (auch die der Alternative X, obwohl die Basishäufigkeit sinkt). Die Versuchspersonen scheinen also auf die Änderung in der Umwelt zu reagieren (und dies, obwohl die hinter dem Experiment liegenden Daten dieselben waren), die zeitlichen Verläufe der guten und schlechten Wasserqualität zu entdecken und entsprechend häufiger auf die Wahl einer Alternative die Wahl derselben Alternative im nächsten Durchgang folgen zu lassen. Diese Tendenz nimmt in T3 weiter zu allerdings nur für die (wesentlich stabilere) Alternative Y.

Tabelle 28: Häufigkeit der Wahl derselben Alternative in der unmittelbar nächsten Runde, aufsummiert über alle 30 Probanden

	T 1 Alternative X	T 1 Alternative Y	T 2 Alternative X	T 2 Alternative Y	T 3 Alternative X	T 3 Alternative Y
Total Wahl	342	162	256	237	258	276
Total Wiederwahl	176	61	166	137	161	198
Anteil	0.51	0.38	0.65	0.58	0.62	0.72

Diskussion

Die doch recht umfassende Darstellung der Ergebnisse verlangt zunächst nach einer Zusammenfassung im Sinne der 17 oben aufgestellten Hypothesen. Wir beginnen mit den „Routineüberprüfungen“ Validitätsfolge und Leistung: Die Häufigkeiten der durchgeführten Tests entsprechen in allen drei Bedingungen (wie nicht anders zu erwarten) der Suchregel in absteigender Validitätsfolge (Hypothese 16). Die Leistungen der Versuchspersonen unterscheiden sich zwischen den drei Bedingungen nicht stark voneinander, nehmen allerdings von 11.8 erfolgreichen Wasserabfüllungen in T1 über 12.6 in T2 bis zu 12.9 in T3 mit zunehmender Anzahl gespielter Durchgänge leicht zu (weil damit die Chancen steigen, zusätzliche gute Wasserlieferungen zu finden). Innerhalb jeder Bedingung erweisen sich die Personen, die mehr Durchgänge mit ihrem Zeitbudget spielen konnten, als die besseren (Korrelationen jeweils $p < .01$ signifikant), so dass wir den gerade bei $p < .05$ signifikanten Unterschied zwischen T1 und T2 ebenfalls diesem Effekt zunehmender Spieldurchgänge zuschreiben können (Hypothese 17).

Für die Hypothesen 1 – 8 ist nicht nur festzustellen, ob sie eintraten oder nicht, sondern es ist gleichermassen interessant zu sehen, ob dafür die reine Einführung der zeitlich dynamischen Umwelt ausreicht oder ob die Versuchspersonen für die erwartete Verhaltensänderung auf das Wissen über den Grad an Stabilität der Zielvariablen angewiesen sind.

Die erste in dieser Hinsicht zu überprüfende Variable ist die Anzahl durchgeführter Tests pro Durchgang (Hypothese 1). Wie erwartet, verringert sich diese Anzahl signifikant, und zwar

bereits von T1 (im Mittel 2.8 Tests) zu T2 (2.3 Tests). Von T2 zu T3 (2.0 Tests) ist dann noch einmal eine signifikante Verringerung festzustellen. Das zusätzliche Wissen um die Stabilität erlaubt den Versuchspersonen also, ihre Tests noch gezielter einzusetzen.

Die Anzahl der Durchgänge, in denen beide Lieferungen weggeschüttet werden (Hypothese 2), steigt gegenüber T1 ($M_{T1} = 9.7$) bereits in T2 massiv an ($M_{T2} = 14.4$), um dann in T3 in etwa auf diesem Niveau zu verbleiben ($M_{T3} = 13.7$). Auch die Anzahl gespielter Runden (Hypothese 3) steigt entsprechend von T1 ($M_{T1} = 26.6$) zu T2 hin massiv an ($M_{T2} = 31.2$) und bleibt dann in T3 auf diesem Level ($M_{T3} = 31.5$). Die Versuchspersonen ergreifen die strategische Möglichkeit des Abwartens, die durch die Zeitachse ermöglicht wird, also bereits unmittelbar. Das Wissen um die Stabilität der Zielvariablen hat hierauf keinen Einfluss.

Mit dem Einsatz dieser Strategie passt sich das Stoppverhalten entsprechend an: War das Entscheiden ganz ohne Information in der statischen Umwelt T1 keine Option ($M_{T1} = 0.3$), so wird dies zu einer, sobald in T2 auf bereits vorliegende Informationen aus der Vergangenheit zurückgegriffen werden kann ($M_{T2} = 4.0$). In T3 ($M_{T3} = 5.2$) steigt dieser Wert noch einmal leicht, aber nicht signifikant an. Auch Hypothese 4 kann also bestätigt werden, und zwar wie Hypothese 2 und 3 bereits im Wechsel von T1 zu T2. Erst später (von T2 zu T3) bestätigt sich allerdings Hypothese 5, die eine Abnahme des *More Reason Decision Making* vorhersagte ($M_{T2} = 12.6$, $M_{T3} = 10.4$; $p < .05$). Es sind die von T1 zu T2 zusätzlich gewonnenen Durchgänge, die – zumindest auf der Ebene der gesamten Stichprobe – in die Erhöhung der Entscheidungen ohne aktuelle Informationssuche gesteckt wird (wie gezeigt). Die MRStR hingegen bleibt in ihrer Häufigkeit von der ersten Bedingung zur zweiten konstant ($M_{T1} = 12.4$).

Differenzierter wird dieses Bild mit einer Betrachtung der individuellen Personen bzw. der Subgruppen, die wir aufgrund des individuellen Stoppverhaltens in T1 gebildet haben: In der Tat sind es die Personen, die in T1 mehrheitlich *One Reason Decision Making* betrieben, welche die Tendenz zeigen, immer einmal wieder ganz ohne neue Informationen zu entscheiden (Hypothese 6, im Resultateteil in Verbindung mit Tabelle 26 im Detail erläutert). Wer hingegen in T1 zum *More Reason Decision Making* neigt, der zeigt die Tendenz, die Cue-Informationen aus der Vergangenheit mit zu berücksichtigen und kommt deswegen in der jeweils aktuellen Informationssuche häufiger einmal mit einer One Reason Stopping Rule aus (Hypothese 7, ebenfalls detailliert in Verbindung mit Tabelle 26 ausgeführt). Allerdings haben wir bei der Besprechung dieser Resultate gesehen, dass nur jeweils ein Teil (12 von 16 Personen von der MRStR zur ORStR sowie 8 von 14 Personen von der ORStR zur NNRStR) dieser Tendenz folgt, einige allerdings so intensiv, dass sich dadurch ihre Majoritätsstrategie verschiebt. Insgesamt reagieren somit 20 der 30 Versuchspersonen (67%) mit einer angepassten Stoppregel auf die Veränderung der Umwelt, ein Drittel reagiert nicht darauf und behält seine Stoppregel bei (bzw. ändert sie im Fall von Vp21 in die nicht erwartete Richtung). Nicht eindeutig auszumachen ist eine Reaktion bereits in T2 oder erst zu T3. Hier vermittelt die Gesamtstichprobe ein uneinheitliches Bild (11 der 20 Personen verändern ihr Stoppverhalten in T2, 9 in T3).

In Hypothese 8 hatten wir uns darauf bezogen, dass im gesamten Experiment hinter den Zielvariablen (und damit auch hinter den Prädiktorvariablen) bestimmte Zeitverläufe stehen, die allerdings in T1 für die Versuchspersonen irrelevant sind (da sie ja unabhängige Durchgänge

spielen) und in T2 noch völlig intransparent. Erst mit den Stabilitätsangaben in T3 erhalten die Probanden eine Ahnung darüber, wie diese Zeitverläufe aussehen können. Allerdings wissen sie dadurch natürlich noch nicht, wie sich diese Verläufe über das Experiment verteilen. Doch das können sie mit der Informationssuche ja herausfinden. Dies sollte, so unsere Erwartung, darin resultieren, dass die Wahrscheinlichkeit dafür steigt, dass dieselbe Alternative, wenn sie beim vorausgehenden Durchgang zum Abfüllen gewählt wurde, unmittelbar wieder gewählt wird. Wegen der starken Asymmetrie der gewählten Alternativen (die wir später noch zu diskutieren haben) sind die Werte in Tabelle 28 nach Alternativen aufgeschlüsselt. Man sieht darin aber sehr deutlich, dass der generelle Anstieg (gleich für welche Alternative) von T1 zu T2 erfolgt: Für Alternative X von 0.51 auf 0.65 und für Alternative Y von 0.38 auf 0.58. Bereits die Einführung einer Zeitachse reicht also aus, damit die Personen mit Hilfe der Informationssuche die zeitlichen Verlaufscharakteristiken der Zielvariablen erkennen. Ein Wissen über den Stabilitätsgrad der Zielvariablen ist dazu nicht notwendig.

Dieses Wissen hat allerdings trotzdem noch Auswirkungen, wie die Überprüfung der Hypothesen 9 – 12 zeigt. Die Informationssuche verlagert sich in T3 im Vergleich zu T2 stärker auf Quelle Y (Hypothese 9): In T2 bestand die noch zu diskutierende Asymmetrie (1113 Tests für Alternative X, 981 Tests für Alternative Y, $p < .05$), in T3 ist dieser signifikante Unterschied dann praktisch vollständig aufgehoben (890 Tests für Alternative X, 860 Tests für Alternative Y, $p = .81$). Dasselbe gilt – zumindest als Tendenz – auch für die Anzahl der Entscheidungen (Hypothese 10): Wie Tabelle 28 zeigt, entscheiden sich in T2 die Personen häufiger für Alternative X (256mal) als für Alternative Y (237mal). Während der Wert für Alternative X in T3 praktisch unverändert bleibt (258mal), wird Alternative Y attraktiver (276mal) und überflügelt Alternative X. Allerdings verfehlt der Anstieg der Attraktivität das Kriterium der Signifikanz ($p = .18$). Bereits beim Aufstellen von Hypothese 10 war freilich kein grosser Effekt erwartet worden, da ja beide Alternativen gleich häufig (ein Drittel der Fälle) eine gute Wasserqualität bieten.

Weitaus deutlicher wird dieser Effekt, wenn man das Verhalten bei direkt zuvor abgefüllter Alternative Y bzw. Alternative X betrachtet: Wenn die stabilere Alternative Y im vorausgehenden Durchgang als positiv bewertet wurde (also zur Abfüllung kam), dann erhöht sich im aktuellen Durchgang in T3 die Wahrscheinlichkeit der Testdurchführung von 41% (in T2) auf 62% (vgl. Tabelle 25). Das bestätigt Hypothese 11. Gleichzeitig steigt auch die Wahrscheinlichkeit der wiederholten Abfüllung an, und zwar von 58% auf 72% (siehe Tabelle 28). Auch Hypothese 12 bestätigt sich also. Ein Anstieg wie bei der stabilen Alternative Y geschieht bei der weniger stabilen Alternative X nicht. Im Gegenteil, die Wahrscheinlichkeit der wiederholten Abfüllung sinkt sogar leicht (von 65% auf 62%), genauso wie die Wahrscheinlichkeit der Informationssuche direkt nach einer Abfüllung (von 79% auf 71%).

Ihre Grenze findet die Verwendung der Stabilitätsinformation allerdings beim Vertrauen auf zeitlich vorausliegende Tests. Die Resultate zu den Hypothesen 13 bis 15 ergeben zwar für T3 alle eine Tendenz zugunsten der stabileren Alternative Y, jedoch wird keiner der t-Tests signifikant: Personen umgehen bei Entscheidungen für Alternative Y nur leicht häufiger den validesten Cue (25mal gegenüber 15mal für Alternative X), sie tendieren nur etwas häufiger dazu, ganz ohne neuen Test für Alternative Y zu entscheiden (46mal gegenüber 35mal für Alternative X), und die 16 More-Reason-Decision-Maker aus T1 vertrauen in T3 nur etwas

häufiger bei einer Entscheidung zugunsten von Alternative Y auf einen einzigen neuen Test (165mal gegenüber 135mal für Alternative X). Die Tendenz geht zwar jedes Mal in die erwartete Richtung, signifikant wird dieser Unterschied aber zu keiner der drei Hypothesen. Wir stellen somit fest, dass die Stabilitätsinformation über die Alternativen das Vertrauen in Information aus der Vergangenheit in diesem Experiment nicht systematisch erhöht.

In der Summe bestätigen die Resultate jedoch fast alle aufgestellten Hypothesen. Die Versuchspersonen reagieren auf die veränderte Umwelt, und zwar in der Regel bereits direkt mit der Einführung einer Zeitachse. Lediglich bei der Anpassung der Stoppregel teilt sich die Stichprobe in solche Personen auf, die unmittelbar in T2 reagieren, in solche, die erst in T3 reagieren, und in solche, welche überhaupt nicht auf die Veränderung der Umwelt reagieren. Dass diese drei Subgruppen praktisch gleich gross sind, sollte aufgrund der begrenzt grossen Stichprobe zunächst einmal als eine Zufälligkeit in den Resultaten betrachtet werden und kann nicht als ein quantitatives Ergebnis dieser Studie herhalten. Was vor allem das letzte Drittel deutlich macht ist, dass die Reaktion auf die Veränderung der Umwelt keine Selbstverständlichkeit ist. Vermutlich decken die signifikanten Mittelwertsunterschiede zu, dass sich auch bei denjenigen Massen, die für die Hypothesentestung nur auf Mittelwertebene geprüft wurden, längst nicht alle Versuchspersonen in die jeweils gefundene Richtung verhalten.

Besonders zu diskutieren ist die enorme Asymmetrie, die sich vor allem in Bedingung T1 im Suchverhalten zwischen Alternative X und Alternative Y zeigt: Für Alternative X werden in der Regel mehr Tests durchgeführt als für Alternative Y, und beim positiven Ausgang wird dann auch Alternative X massiv häufiger gewählt. Diese Asymmetrie dürfte auf die Gewohnheit zurückzuführen sein, auf einem Bildschirm (oder auf einem Blatt Papier) oben links zu beginnen (sofern es keinen Grund gibt, an anderer Stelle einzusteigen). Und links steht eben immer Alternative X. Insofern ist diese Asymmetrie für die Beantwortung der Fragestellungen irrelevant. Ja sie stört nicht einmal wirklich, denn dadurch wird – unvorhergesehenerweise – in Bedingung 1 eine Befundlage geschaffen, die sich in Bedingung 2 und 3 replizieren sollte, wenn die Versuchspersonen keinen markanten Unterschied in den Bedingungen sehen und damit keinen Grund, an anderer Stelle als oben links mit einem jeden Durchgang zu beginnen. Die Asymmetrie reduziert sich aber in Bedingung 2 markant und verschwindet ganz in Bedingung 3. Das ist also ein zusätzlicher Hinweis darauf, dass die Versuchspersonen auf die geänderte Umwelt eingehen.

Schlussbemerkung

Die Resultate zeigen, dass Personen in dem experimentellen Setting ihr Verhalten eindeutig an die neu eingeführte Zeitachse anpassen. Zusammenfassend kann gesagt werden, dass sie die Informationssuche im zeitlichen Kontext reduzieren, sich ab und zu auch ohne eine aktuelle Information entscheiden und häufiger auch das „Nichtstun“ wählen. Die Personen beziehen somit vergangene Informationen mit in die Planung der anstehenden Informationssuche ein und optimieren so ihr Suchverhalten in Bezug auf das Kosten-Nutzen-Verhältnis. Die mit der Zeitdimension einher gehende Stabilität wird zum Zeitsparen genutzt und erbringt damit einen prospektiven Gewinn.

Wird eine Zeitachse eingeführt, treten zudem vermehrt Wiederholungen bei der Suche und bei der Wahl einer Alternative auf. Dieses Verhalten ist sinnvoll, da in der zweiten Bedingung T2 immer dieselben beiden Quellen zur Auswahl stehen. Dies verleitet die Versuchspersonen möglicherweise zur Annahme, dass die Zielvariablen über eine gewisse Stabilität verfügen. Sie liegen damit auch richtig, müssen jedoch das Pattern durch aktive Informationssuche entdecken.

Das Experimentaldesign hat somit zu einer Reihe von ersten Erkenntnissen zur Implementierung einer Zeitachse und zur Implementierung der Ressource Zeit im Informationssuch- und Entscheidungsverhalten geführt. Gleichzeitig war das Design jedoch in der Kosten-Nutzen-Struktur etwas asymmetrisch aufgebaut, indem das Wegschütten von Wasser keinerlei Ressourcen benötigte. Versuchspersonen hätten also ab T2 auf die Strategie kommen können, bei negativen Testinformationen zu beiden Wasserlieferungen einfach ohne weitere Tests eine ganze Reihe von Wasserlieferungen einfach wegzuschütten und erst dann wieder zu testen, ob sich die Situation gebessert hätte. Diese Strategie (welche lediglich von Vp14 und von Vp25 gehäuft angewendet wurde), ist ein Zeichen der vollständigen Einsicht in die statistische Struktur der Umwelt, jedoch „verwässert“ sie gleichzeitig die Auswertung im Hinblick auf die meisten der 17 operationalisierten Hypothesen. Insofern waren wir froh, dass lediglich zwei Probanden diese Strategie verwendet haben. Für zukünftige Experimente mit diesem Trinkwasser-Design würden wir jedoch überlegen, ob es nicht günstiger wäre, auch für das Wegschütten einer Wasserladung jeweils einen Aufwand (z.B. 1 Stunde pro Lieferung) anzusetzen.

Ob die Festlegung auf eine fixe Reihenfolge der drei Experimentalbedingungen eine Limitation für die Resultate darstellt oder nicht, kann kontrovers diskutiert werden. Auf der einen Seite ist es guter Brauch, im Falle mehrerer Bedingungen die Reihenfolge ausgewogen oder mindestens doch randomisiert zu halten. Das setzt jedoch eine gewisse Austauschbarkeit der Bedingungen voraus. Hinsichtlich des hier eingeführten Szenarios erscheint uns diese Austauschbarkeit jedoch nicht gegeben zu sein: Wenn einmal die Zeitachse (T2) eingeführt ist, dann ist fraglich, ob jede Versuchsperson diese auch wieder ausblenden kann, sollte sie erst danach mit T1 konfrontiert werden. Und noch kritischer erscheint uns die Aufgabe hinsichtlich einer gekehrten Reihenfolge T3 / T2 zu sein: Wenn man den Versuchspersonen einmal die Stabilität der Quellen X und Y bekannt gegeben hat, wie garantiert man dann, dass sie diese Information danach ausblenden? Allein die Umbenennung in zwei völlig neue Quellen würde diesen Effekt nicht zwingend erzielen, da es ja nicht um die Abstraktion von den konkreten Quellen geht, sondern um die Tatsache, sich zuvor bereits in einer statistischen Umwelt mit expliziter Stabilitätsinformation bewegt zu haben und somit für die Frage der Stabilität sensitiviert worden zu sein. Diese Sensitivierung (so es einer solchen bedarf) ist, einmal geschehen, nicht wieder rückgängig zu machen.

Eine mögliche Lösung für dieses Problem bestünde darin, dass man auf abhängige Daten verzichtet und jede Versuchsperson nur eine der drei Bedingungen spielen lässt. Allerdings erscheint uns gerade der within-Vergleich sehr aussagekräftig zu sein, einmal weil er etwas über das Verständnis der Versuchspersonen hinsichtlich der Veränderung der Umwelten aussagt und zum anderen, weil er für T2 und T3 Aussagen darüber ermöglicht, ob One-Reason-Decision-Maker auf die Einführung der Zeitachse und der Stabilitätsinformation anders rea-

gieren als More-Reason-Decision-Maker. Diese grundsätzlich unterschiedlichen Strategien sind in der bislang vorliegenden Literatur anhand repetitiver Einzeldurchgänge gut eingeführt (vgl. die in der Einleitung umfassend zitierten Arbeiten), so dass es sich anbietet, diese Personen anhand der Bedingung T1 zu identifizieren und dann zu schauen, wie weit sich die Veränderungen in der statistischen Umwelt auf die beiden grundsätzlichen Strategien auswirken.

Ohnehin erscheint es uns aussichtsreich, bei zukünftigen Experimenten in Design und Auswertung stärker auf diese Distinktion der Strategien (frugal versus absichernd) zu setzen. Auch im Trinkwasserszenario konnten diese Strategien als verhältnismässig stabil identifiziert werden (s. noch einmal Tabelle 26). Je komplexer die experimentell untersuchten Umwelten werden, umso variantenreicher wird auch die Verhaltensanpassung einer Person an die Veränderungen ausfallen. Und vor diesem Hintergrund scheint es ein guter Ausgangspunkt für die Auswertungen zu sein, wenn man etwas über die grundlegende Einstellung einer Person zur Stopregel weiss.

Die Einführung der Zeitachse ist eine solche Erhöhung des Komplexitätsgrads. Die Versuchspersonen zeigen, dass sie damit umzugehen in der Lage sind, und zwar vornehmlich im Hinblick auf die Organisation ihres künftigen Verhaltens: Vor allem aus der zurückliegenden Test-Information und aus dem der Stabilitätsgrad der Wasserqualität der Quellen schliessen sie, bei welcher Alternative es sich lohnt weiterzumachen. Das macht ihr Verhalten auf Stichprobenebene effizienter (ohne dass diese Effekte allerdings so gross wären, dass sie bei jedem Individuum klar ablesbar wären. Dazu müssten vermutlich die Experimentaldurchgänge viel länger sein.) Bemerkenswert ist, dass sie dafür kein quantitatives Wissen über die Stabilität der Alternativen benötigen: Allein die Einführung der Zeitachse führt bereits zu den massgeblichen Verhaltensänderungen.

Nicht oder praktisch nicht genutzt wird die Stabilitätsinformation hingegen im Hinblick auf die Nutzung vergangener Information. Hier könnten die Versuchspersonen in einer wesentlich stärkeren Weise mutmassen, die ein Testresultat aus dem letzten Durchgang auch noch bei der aktuellen Wasserlieferung bestand hat, wenn die Stabilität so hoch ist wie bei Alternative Y (.90 nämlich). Dies geschieht jedoch nicht. Es wird deswegen für zukünftige Forschung spannend sein, eine grundsätzliche Asymmetrie der Informationsnutzung im Hinblick auf Vergangenheitsnutzung und Zukunftsplanung zu überprüfen und, sofern sie sich erhärten sollte, ihren psychologischen Gründen nachzuforschen.

Literatur

- Bröder, A. (2000a). Assessing the empirical validity of the "Take The Best"-heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1332-1346.
- Bröder, A. (2000b). "Take the best - Ignore the rest" Wann entscheiden Menschen begrenzt rational? Lengerich: Pabst.
- Bröder, A. (2002). Take The Best, Dawes' Rule, and compensatory decision strategies: A regression-based classification method. *Quality & Quantity*, 36, 219-238.
- Bröder, A. (2003). Decision making with the "adaptive toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 611-625.
- Bröder, A. (2005). Entscheiden mit der "adaptiven Werkzeugkiste". Ein empirisches Forschungsprogramm. Lengerich: Pabst Science Publishers.

- Bröder, A., & Schiffer, S. (2003). Take The Best versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General*, 132, 277-293.
- Bröder, A., & Schiffer, S. (2006). Adaptive flexibility and maladaptive routines in selecting fast and frugal decision strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(4), 904-918.
- Christen, S., Zurbruggen, S., Hausmann, D., & Läge, D. (2007). Trinkwasser- und Börsenszenario: Aufbau, Ablauf und technische Spezifikation einer Experimentalumgebung zur entscheidungsvorbereitenden Informationssuche in zeitlichen Verläufen. AKZ-Forschungsbericht Nr. 43. Zürich: Angewandte Kognitionspsychologie.
- Gigerenzer, G., & Goldstein, D. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650-669.
- Gigerenzer, G., Todd, P. M., & Gerd Gigerenzer, A. R. (Eds.). (1999). *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.
- Hausmann, D., Christen, S., & Läge, D. (2006). Urteils-Sicherheit um jeden Preis? Verletzung der Stopppregel in Abhängigkeit von der Höhe der Informationssuchkosten In B. Gula, R. Alexandrowicz, S. Strauß, E. Brunner, B. Jenull-Schiefer & O. Vitouch (Eds.), *Perspektiven psychologischer Forschung in Österreich. Proceedings zur 7. Wissenschaftlichen Tagung der Österreichischen Gesellschaft für Psychologie* (pp. 45-51). Lengerich: Pabst.
- Hausmann, D., & Läge, D. (2005). Wenn Limiten zum Strategiewechsel führen: Wechsel der Suchstrategie unter situativen Einschränkungen. AKZ-Forschungsbericht Nr. 6. Zürich: Angewandte Kognitionspsychologie.
- Hoffrage, U., Hertwig, R., & Czienskowski, U. (2003). The ecological rationality of Take The Best's building blocks. Vortrag an der 19th bi-annual conference on subjective Probability Utility and Decision Making (SPUDM). Zurich: Swiss Federal Institute of Technology.
- Kerstholt, J. H. (1994). The effect of time pressure on decision - making behavior in a dynamic task environment. *Acta Psychologica*, 86, 89-104.
- Läge, D., Hausmann, D., & Christen, S. (2005). Wie viel bezahlen für eine valide Information? Suchkosten als limitierender Faktor der Informationssuche. AKZ-Forschungsbericht Nr. 7. Zürich: Angewandte Kognitionspsychologie.
- Läge, D., Hausmann, D., Christen, S., & Daub, S. (2005). Was macht einen "guten Cue" aus? Strategien der Informationssuche beim heuristischen Entscheiden unter Unsicherheit. AKZ-Forschungsbericht Nr. 5. Zürich: Angewandte Kognitionspsychologie.
- Newell, B., Rakow, T., Weston, N., & Shanks, D. (2004). Search strategies in decision making: The success of 'success'. *Journal of Behavioral Decision Making*, 17, 117-137.
- Newell, B., & Shanks, D. (2003). Take the best or look at the rest ? Factors influencing one-reason decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 53-65.
- Newell, B., Weston, N., & Shanks, D. (2003). Empirical tests of a fast and frugal heuristic : not everyone "takes-the-best". *Organizational Behavior and Human Decision Processes*, 91, 82-96.
- Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, 16, 366-387.
- Payne, J. W., & Bettman, J. R. (2000). Preferential Choice and Adaptive Strategy Use. In G. Gigerenzer & R. Selten (Eds.), *Bounded Rationality: The Adaptive Toolbox*. Cambridge, London: MIT Press.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 534-552.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press.
- Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics and how can we tell? In G. Gigerenzer, P. Todd & A. R. Group (Eds.), *Simple Heuristics That Make Us Smart* (pp. 141-168). New York: Oxford University Press.
- Rieskamp, J., & Hoffrage, U. (2006). Inferences under time pressure: How opportunity costs affect strategy selection. . Submitted for publication.
- Zurbruggen, S., Christen, S., Hausmann, D., & Läge, D. (2007a). Adaptivität der Informationssuche in statischen und in dynamischen Umwelten. AKZ-Forschungsbericht Nr. 46. Zürich: Angewandte Kognitionspsychologie.
- Zurbruggen, S., Christen, S., Hausmann, D., & Läge, D. (2007b). Informationssuche beim Entscheiden in zeitlich dynamischen Umwelten. AKZ-Forschungsbericht Nr. 44. Zürich: Angewandte Kognitionspsychologie.
- Zurbruggen, S., & Läge, D. (2007). Ressource Zeit statt Geld: Ein Vergleich zweier Szenarien zur Entscheidungsforschung. AKZ-Forschungsbericht 45. Zürich: Angewandte Kognitionspsychologie.

Kapitel 5: Ressource Zeit statt Geld: Ein Vergleich zweier Szenarien zur Entscheidungsforschung

Einleitung

Wir leben in einer Welt, in der Zeit – zumindest in unseren Breitengraden – als eine der wertvollsten Ressourcen betrachtet wird. Gerade aus dem Blickwinkel einer ökologischen Rationalität ist es daher sinnvoll, diese Ressource auch in psychologischen Entscheidungsmodellen mit einzubeziehen. Bis anhin wurden die Kosten, welche bei der Informationsbeschaffung vor Entscheidungen anfallen, gemeinsam mit den möglichen Gewinnen nach einer richtigen Entscheidung (also dem Nutzen) mit der Ressource Geld experimentell implementiert. (Diese Operationalisierung wird historisch und sachlich vor dem ökonomischen Hintergrund nachvollziehbar, vor dem sich Theoriebildung und empirische Forschung im Bereich der Entscheidungspsychologie entwickelt haben). Wir erachten es jedoch als sinnvoll, zusätzlich dazu auch die Ressource Zeit in Entscheidungsmodelle einzuführen. Diese Ressource zeichnet sich im Erleben eines Individuums dadurch aus, dass sie kontinuierlich verbraucht wird. Man kann sie also weder ansparen (indem man eine Zeit lang nichts tut) noch in beliebiger Häufung einsetzen (wie etwa Geld). Ein weiterer struktureller Unterschied zu einer monetären Ressource besteht darin, dass Zeit nicht die unmittelbare „Währung“ des Erfolgs einer Entscheidung ist (zumindest nicht im Regelfall), so dass der Nutzen nicht unmittelbar mit dem Aufwand in einer Kosten-Nutzen-Analyse verrechnet werden kann.

Trotz dieser markanten Unterschiede ist jedoch auch die Ressource Zeit quantifizierbar, und das heisst kompetitiv zueinander in Beziehung setzbar (man kann mehr Zeit für eine Handlung einsetzen oder weniger). Für die Möglichkeit einer Nutzensmaximierung ist das eine unabdingbare Voraussetzung. Bevor man nun in detaillierte Experimente und in eine psychologische Theoriebildung über den Umgang mit dieser Ressource in Entscheidungssituationen einsteigt, ist deswegen sinnvoll zu wissen, ob Menschen die Ressource Zeit fundamental anders einsetzen als die Ressource Geld oder ob die erfüllte Grundbedingung der Quantifizierbarkeit ausreicht, um die bei der Ressource Geld bereits bekannten Verhaltensmuster zu evozieren.

Im Folgenden wird deswegen ein explorativ ausgerichtetes Experiment berichtet, welches in einem Trinkwasserszenario die Ressource Zeit in Parallelität zu einem monetär ausgerichteten Börsenszenario verwendet. (Beide Szenarien sind in ihrem detaillierten Aufbau und ihrer technischen Spezifikation bei (Christen, Zurbruggen, Hausmann, & Läge, 2006) ausführlich beschrieben). Kurz zusammengefasst geht es für die Versuchsperson darum, mit einem gesamthaften Zeitbudget von 240 Stunden Spielzeit möglichst viel geniessbares Trinkwasser in Flaschen abzufüllen. Jede eingehende Wasserlieferung kann sie mit bis zu vier probabilistischen Tests auf die Güte hin testen, wobei sie für jede Testdurchführung eine Stunde benötigt. Entscheidet sie sich dafür, das Wasser aus einer der beiden Lieferungen abzufüllen, so ist sie zehn Stunden damit beschäftigt. Erst zu einem späteren Zeitpunkt erfährt sie, ob die abgefüllten Flaschen die staatliche Kontrolle überstanden haben oder ob sich das Wasser doch als ungeniessbar herausgestellt hat. Die Versuchsperson erhält parallel jeweils zwei Wasserlieferungen aus unterschiedlichen Quellen, wobei sie sich (nach allfälliger Durchführung von

Tests) entweder für eine der beiden Lieferungen entscheiden kann oder aber auch beide Lieferungen wegschütten lassen kann.

Damit ist dieses Experiment parallel aufgebaut zum bereits angesprochenen Börsenszenario, in welchem insgesamt 20 Spielrunden auf einem Information Board (vgl. Payne, 1976) absolviert werden. In jeder Runde stehen dort der Versuchsperson 10'000 Schweizer Franken zur Verfügung. Diese kann sie entweder auf einen Optionsschein setzen (und damit im Gewinnfall den Betrag verdreifachen), wobei die Option in einem Drittel der Fälle steigend in ihrem Wert ist. Oder das Geld kann direkt aufs Gewinnkonto transferiert werden, wobei man nichts zusätzlich gewinnen, aber auch nichts verlieren kann. Bevor die Person sich entscheidet, kann sie Informationen von vier verschiedenen probabilistischen Indikatoren (Cues) zu den zwei zur Verfügung stehenden Optionen kaufen. Die vier Indikatoren geben mit einer gewissen Trefferrate (der Validität, so wie z.B. bei Gigerenzer & Goldstein (1996) eingeführt) an, ob ein Optionsschein an Wert gewinnen oder zum Verlust des Kapitals führen wird. Die Versuchspersonen können die Cues in freier Kombination, aber immer streng sequentiell aufdecken. Die Validitäten sind in 5%-Schritten abgestuft (.90, .85, .80, .75) und werden den Versuchspersonen auf dem Information Board als Trefferraten präsentiert.

Informationen

Indikator	Trefferrate	Optionsschein X	Optionsschein Y
Indikator A	80	<input type="checkbox"/>	<input type="checkbox"/>
Indikator B	75	<input type="checkbox"/>	<input type="checkbox"/>
Indikator C	90	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Indikator D	85	<input type="checkbox"/>	<input type="checkbox"/>

Handel

Budget 10'000 minus Suchkosten 2'000 = investierbarer Betrag 8'000

Optionsschein X: 0
Optionsschein Y: 0
Sparbuch: 8'000

Auswahl bestätigen

Übersicht

Bisheriger Kontostand: 275'000
Gewinn in den letzten 5 Tagen: 89'000 (darunter 1 Tage ohne Gewinn)
Neuer Kontostand: 364'000

Anzahl vergangener Handelstage: 19

Weiter

Abb.16: Informationboard des Börsenszenarios

Analyse

Test	Trefferrate	Quelle X	Quelle Y
Test A	75	<input type="checkbox"/>	<input type="checkbox"/>
Test B	80	<input type="checkbox"/>	<input type="checkbox"/>
Test C	90	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Test D	85	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Legende: ☒ = schlechte Wasserqualität ☒ = gute Wasserqualität

Handlung

Wasser aus Quelle X: ☐
Wasser aus Quelle Y: ☒
Keine Abkühlung: ☐

Auswahl bestätigen

Übersicht

Aktueller Zeitverbrauch (in Stunden): 13
Noch vorhandene Zeit (in Stunden): 215

Anzahl brauchbare Lieferungen bisher: 0
Von den letzten 1 Lieferungen brauchbar: 0
Anzahl brauchbare Lieferungen insgesamt: 0

Nächste Lieferungen

Abb.17: Informationboard des Trinkwasser-Szenarios

Anhand dieses Settings kann nun die Informationssuche und das Stoppverhalten der Versuchspersonen vor einer Entscheidung untersucht werden. Im Börsenszenario hatte sich – in Übereinstimmung mit der bereits vorliegenden Forschung in diesem Feld – gezeigt (siehe detaillierte Resultate bei Zurbruggen, Christen, Hausmann, & Läge (2007)), dass die Versuchspersonen die Informationssuche durchgängig an den Validitäten der Informationen orientieren (Suchregel). Bei der Beendigung der Informationssuche zeigt sich der in dieser Form von Experimenten übliche Mix zwischen *One Reason Decision Making* (d.h. dem Stoppen, sobald ein positiver Hinweis vorliegt, siehe auch die Definition von ORDM bei Gigerenzer & Goldstein (1996)) und *More-Reason Decision Making* (also dem Absichern der vorliegenden Information durch einen oder mehrere weitere Cues, siehe dazu Definition von Hausmann & Läge (2005)). In ihrer Entscheidung folgen die Versuchspersonen der Richtung von dem oder den Cues und zeigen sich dabei durchaus risikobewusst. Dieses Börsenszenario evoziert somit

das aus anderen Experimenten bereits bekannte Verhaltensmuster und erscheint uns deswegen als eine solide Grundlage für den Vergleich mit der Ressource Zeit.

Uns interessiert nun, wie die Personen mit der Ressource Zeit umgehen und ob sich ein Unterschied im Verhalten zum Börsenszenario zeigt. Wir implementieren Zeit jedoch nicht real als Faktor des Experimentes, wie dies u.a. in der Forschung zu komplexem Problemlösen, dynamic decision making und naturalistic decision making der Fall ist. Dort fand man beispielsweise, dass Personen häufig zu lange nach Informationen (wie z.B. Testresultaten) suchen, statt schneller zu handeln, um ein optimales Resultat zu erreichen (Kleinmuntz and Thomas 1987; You 1989; Kerstholt and Raaijmakers 1997; Huber and Macho 2001). Im hier zu berichtenden Experiment ersetzen wir vielmehr die Ressource Geld durch fiktive Zeit im Bezug auf die Informationskosten und den zum Erzielen des Gewinns zu erbringenden Aufwand: Statt einen Cue für ein Zehntel des Budgets zu kaufen, führen die Versuchspersonen einen Test durch, für den sie eine Stunde vom fiktiven Zeitbudget benötigen. Und statt Geld auf eine Option zu setzen, verbrauchen sie 10 Stunden Zeit, um das für gut befundene Trinkwasser in Flaschen abzufüllen. Ansonsten sind die Szenarien, inklusive der Cues und ihrer Validitäten, parallel zueinander aufgebaut.

Für den Vergleich der beiden Ressourcen sind nun folgende Werte (im Sinne abhängiger Variablen) von Interesse:

1. Informationssuchverhalten (angeschaute Indikatoren): Bei der Ressource Geld orientieren sich Personen an der Validitätsfolge. Das sollte bei der Ressource Zeit nicht anders sein. Im gewählten Szenario werden allerdings die Cue-Informationen getrennt für beide Alternativen erhoben, so dass die Versuchspersonen nach dem Befragen des validesten Indikators für eine der beiden Alternativen entweder mit dem zweitvalidesten Indikator für dieselbe Alternative oder mit dem validesten Indikator für die andere Alternative fortfahren können. Selbst bei der (trivialen) Beachtung der Validitätsfolge können sich für die einzelnen Cues somit unterschiedliche Häufigkeiten ergeben.

2. Beenden der Informationssuche: In der Tradition der Forschung zum Take the Best-Modell (Gigerenzer & Goldstein, 1996) als prominentestem Vertreter des *One Reason Decision Making* lassen sich drei Stoppregeln voneinander unterscheiden: a) Raten, also Entscheiden ohne Hinzuziehen eines Cues, b) One-Reason Stopping Rule (ORStR), also Abbruch der Informationssuche, sobald ein für die Entscheidung nutzbarer Cue vorliegt (im gewählten Szenario wäre das ein positiver Indikator für eine der beiden Alternativen oder je ein negativer für beide Alternativen), c) More-Reason Stopping Rule (MRStR), das Weitersuchen nach Information, obwohl Bedingung (b) bereits erfüllt ist. Es ist davon auszugehen, dass in einem solchen Szenario nur wenige Personen substantiell häufig raten werden, also mehrheitlich Strategie (a) verfolgen. Zu überprüfen ist deswegen zweierlei: 2.a) Unterscheidet sich der Anteil von Personen, der mehrheitlich einer ORStR bzw. einer MRStR folgt, in Abhängigkeit von der Ressource? Und 2.b) Führt die Ressource Zeit zu einer klareren (respektive weniger klaren) Ausprägung einer intraindividuellen Stabilität der gewählten Stoppregel als die Ressource Geld?

3. Entscheidungsverhalten: Dass die Personen der durch die Cues angezeigten Richtung folgen werden, ist als Entscheidungsregel im *One Reason Decision Making* trivial. Interessant

ist vielmehr das Ausmass des Risiko-Verhaltens, welches sie a) insgesamt zeigen (Anzahl der Entscheidungen zu investieren bzw. Wasser abzufüllen) und b) in denjenigen Fällen zeigen, in denen sie auf konfligierende Cue-Informationen treffen. Konfligierende Cue-Informationen liegen dann vor, wenn zwei oder mehrere Cues bei einer Alternative aufgedeckt wurden und sowohl Hinweise auf gute wie auch auf schlechte Wasserqualität geben. Hier könnte sich die Ressource Zeit systematisch von der Ressource Geld unterscheiden.

Die Hypothesen sind bewusst zweiseitig formuliert. Da uns bislang keinerlei experimentelle Befunde zum Vergleich dieser beiden Ressourcen bekannt sind, besitzt das zu berichtende Experiment von vornherein einen gewissen explorativen Charakter. Einer vorgängigen Spekulation über die Richtung eines etwaigen Unterschieds enthalten wir uns deswegen und werden – aus demselben Grund – deswegen auch etwaige Unterschiede vorsichtig einzuordnen haben.

Methode

Stichprobe

Insgesamt absolvierten 60 Versuchspersonen (Alter $M = 30.3$, $SD = 8.9$; 30 Männer und 30 Frauen) sowohl das Trinkwasserszenario (Ressource Zeit) als auch das Börsenszenario (mit der klassischen Ressource Geld als Vergleich), und zwar aus Gründen einer einheitlichen Instruktion jeweils in dieser Reihenfolge. Beide Szenarien, die oben bereits skizziert sind, finden sich ausführlich bei (Christen et al., 2006) dargestellt.

Versuchsdurchführung

Während das Börsenszenario aus 20 Durchgängen mit einer Ressource von jeweils 10'000 Franken bestand (die es zu vermehren galt), wurde den Versuchspersonen im Trinkwasserszenario ein Budget von 240 Arbeitsstunden zur Verfügung gestellt. Diese strukturelle Unterscheidung trägt dem Umstand Rechnung, dass Zeit eben nicht wie Geld beliebig quantitativ einsetzbar ist, sondern kontinuierlich vergeht. Das Durchführen eines Tests an einer der beiden zeitgleich zur Verfügung stehenden Wasserlieferungen benötigt eine Arbeitsstunde, das Abfüllen des Wassers 10 Arbeitsstunden. Der Erfolg einer Versuchsperson wird daran gemessen, wie viele qualitativ gute Wasserfüllungen sie schafft, bevor die 240 Arbeitsstunden aufgebraucht sind.

Im Gegensatz zum Börsenszenario (welches aus 20 Durchgängen besteht) ist die Anzahl der Durchgänge im Trinkwasserszenario abhängig davon, wie viele Tests die jeweilige Versuchsperson pro Wasserlieferung durchführt und wie häufig sie sich dafür entscheidet, Wasser abzufüllen. Um eine Vergleichbarkeit zum Börsenszenario zu gewährleisten, werden wo notwendig deswegen nur die ersten 20 Durchgänge für die Auswertung der Daten des Trinkwasserszenarios berücksichtigt.

Resultate

Bei den folgenden inferenzstatistischen Auswertungen (t-Tests für abhängige Stichproben) wurde $\alpha = 0.05$ mit zweiseitiger Hypothesenprüfung als Grundlage für das Akzeptieren eines Mittelwertsunterschiedes gewählt.

Leistung

Die Leistung, d.h. der Gewinn, den die Versuchspersonen in einem Szenario erwirtschaften, beträgt im Trinkwasserszenario im Mittel 12 Punkte (SD= 2.4), im Börsenszenario CHF 294'800 Franken (SD= 21'411). Die Leistungen korrelieren mit $r = .45$ signifikant ($p < .01$); Versuchspersonen, die im Umgang mit der Ressource Geld gut wirtschaften, erweisen sich also in der Tendenz auch als etwas besser im Umgang mit der Ressource Zeit.

Informationssuche

In beiden Szenarien entspricht die Reihenfolge der Cuesuch-Häufigkeit – wie nicht anders zu erwarten – der Rangreihe der absteigenden Validitätsfolge der Cues (Tabelle 29). Die Häufigkeit des validesten Cues unterscheidet sich zwischen den beiden Szenarien nicht signifikant ($p = .64$), die übrigen drei Cues jedoch werden signifikant häufiger im Trinkwasserszenario konsultiert als im Börsenszenario (Cue 2 $p < .01$, Cue 3 $p < .01$, Cue 4 $p < .05$). Insgesamt werden im Trinkwasserszenario auf diese Weise signifikant mehr Informationen gesucht als im Börsenszenario ($p < .01$). Das knappe Viertel, um das sich das Heranziehen der Indikatoren erhöht, geht einher mit einer deutlich grösseren Standardabweichung: Im Trinkwasserszenario verwenden also nicht alle Personen in gleicher Weise mehr Informationen, sondern der Effekt ist auf einen Teil der Stichprobe zurückzuführen.

Tabelle 29: Mittelwert und Standardabweichung der Häufigkeit der Cuesuche bei den beiden Alternativen über die ersten 20 Runden.

Summe von	Trinkwasser	Börse
Cue 1 (V = .90)	M = 29.0, SD = 7.7	M = 28.6, SD = 8.2
Cue 2 (V = .85)	M = 11.4, SD = 11.4	M = 6.3, SD = 7.2
Cue 3 (V = .80)	M = 5.5, SD = 6.6	M = 3.2, SD = 4.6
Cue 4 (V = .75)	M = 4.0, SD = 6.6	M = 2.4, SD = 4.4
Alle Cues	M = 49.9, SD= 23.9	M = 40.6, SD= 14.0

Stoppregel

Gemäss der bei den Hypothesen eingeführten Unterscheidung klassifizieren wir das Stoppverhalten in drei Kategorien: Raten (d.h. Entscheiden ohne Hinzuziehen von Indikatoren), One-Reason Stopping Rule (Entscheidung auf der Basis einer einzigen vorliegenden Information) und More-Reason Stopping Rule (Entscheidung auf der Basis von mehr als einer vorliegenden Cue Ausprägung). Da für jede Alternative die Cues getrennt befragt werden konnten, subsummieren wir unter der More-Reason Stopping Rule (MRStR) alle Fälle, in denen entweder zu einer beliebigen Alternative mehr als ein Cue befragt wurde oder in denen nach Vorliegen einer positiven Cue-Ausprägung bei einer Alternative noch mindestens ein Cue bei der anderen Alternative befragt wurde. Die One-Reason Stopping Rule (ORStR) umfasst entsprechend neben dem Stoppen nach dem ersten Befragten Cue auch diejenigen Fälle, in denen nach einer ersten negativen Cue-Ausprägung für die eine Alternative noch ein weiterer Cue für die andere Alternative befragt wurde.

Tabelle 30 zeigt das Stoppverhalten auf individueller Ebene. Nach Majoritätsregel können im Trinkwasserszenario 30 Personen (50%) als ORStR-Anwender klassifiziert werden und 26 Personen (43%) stoppen überzufällig häufig gemäss einer MRStR. Im Börsenszenario stoppen 38 Personen (63%) häufig nach einer Information (ORStR) und 17 Personen (28%) benötigen

mehrere Informationen (MRStR). In beiden Szenarien entscheidet sich also die Mehrzahl der Versuchspersonen in der Regel für eine ORStR, wobei diese Asymmetrie im Falle der Ressource Zeit jedoch geringer ausfällt als bei der Ressource Geld.

Das verringerte Vorkommen der ORStR bei der Ressource Geld zeigt sich auch auf Mittelwertebene: Im Trinkwasserszenario ist häufiger ($p < .01$) die MRStR ($M = 8.7$, $SD = 7.7$) zu beobachten als im Börsenszenario ($M = 6.4$, $SD = 6.2$). Umgekehrt wird bei der Ressource Zeit seltener ($p < .01$) auf der Basis einer einzigen Information gestoppt ($M = 10.5$, $SD = 7.6$) als bei der Ressource Geld ($M = 13.1$, $SD = 5.9$). Die Ratestrategie ist zu vernachlässigen, da nur die drei Personen 38, 43 und 55 im Trinkwasserszenario häufig raten. (Dies führt jedoch nicht zu einem signifikanten Unterschied zwischen den beiden Bedingungen).

Tabelle 30: Stoppverhalten der 60 Versuchspersonen im Trinkwasser-Szenario (T) und im Börsenszenario (B): Anzahl Runden der Versuchspersonen, in denen sie ohne Befragung eines Cues entschieden haben (0), eine One-Reason Stopping Rule verfolgt haben (1) oder eine More-Reason Stopping Rule einsetzten (>1). Fett markiert ist jeweils die Majoritätsstrategie (mehr als 10 Runden) dieser Versuchsperson. Die Tabelle ist sortiert nach absteigender Anzahl Fälle der ORStR in T und B.

nach absteigender Primzahl-Potenz der Ordnung 1 und 2:																															
	Vp	33	39	41	46	58	22	31	52	53	10	14	23	8	25	3	5	56	54	51	29	32	4	11	24	17	42	27	37	13	40
T	0	0	0	0	0	0	0	0	0	1	0	2	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1	0	0
T	1	20	20	20	20	20	20	20	19	19	18	18	19	17	18	17	20	18	17	20	20	11	15	20	12	8	20	11	13	5	9
T	>1	0	0	0	0	0	0	0	1	0	2	0	0	3	2	3	0	0	3	0	0	9	5	0	8	12	0	9	6	15	11
B	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	2	5	0	0	0	0	0	0	10	1	0	0	0	0	0	0
B	1	20	20	20	20	20	19	19	20	20	20	20	19	20	19	18	15	16	16	12	11	20	15	10	16	19	7	15	13	20	16
B	>1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	4	4	8	9	0	5	0	3	1	13	5	7	0	4
	Vp	48	6	36	43	55	47	50	12	38	18	21	60	30	16	44	34	45	2	1	15	9	19	26	20	49	57	59	28	7	35
T	0	0	0	11	16	0	0	0	6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	
T	1	13	7	5	9	4	11	10	5	9	3	11	6	0	7	7	7	4	1	3	1	1	0	2	0	0	0	0	0	0	1
T	>1	7	13	15	0	0	9	10	15	5	16	9	14	20	13	13	13	16	19	17	19	19	20	18	19	19	20	20	16	20	19
B	0	0	0	1	5	0	0	0	2	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
B	1	12	17	19	15	20	11	11	14	10	14	6	10	15	7	7	6	9	10	6	8	7	6	4	5	5	4	4	3	3	2
B	>1	8	3	0	0	0	9	9	4	10	6	14	10	4	13	13	14	10	10	14	12	13	14	16	15	15	16	16	17	17	18

70% der Versuchspersonen haben über beide Szenarien hinweg eine einheitliche Majoritätsstrategie: 42 Probanden verwenden jeweils dieselbe dominante Stoppstrategie, davon 27 Personen die ORStR und 15 Personen die MRStR. Acht Personen, die im Trinkwasserszenario mehrheitlich einer MRStR folgen, wechseln im Börsenszenario zu einer ORStR. Umgekehrt sind es nur zwei Personen, die von der ORStR im Trinkwasserszenario zur MRStR im Börsenszenario wechseln. Auch dieser Befund geht einher mit dem Gesamtbild, dass die Ressource Zeit mit einer Tendenz zu mehr More-Reason Decision Making einhergeht.

Eine zweite Facette der Auswertung der Stoppregel ist die Überprüfung der Deutlichkeit, mit der das Szenario jede Versuchsperson auf eine der beiden Stoppregeln festlegt. Da bietet sich der Differenzwert zwischen der Fallzahl der ORStR oder der MRStR an. Je grösser dieser Wert, umso deutlicher legt das Szenario die Personen auf eine Strategie fest. Für das Trinkwasserszenario erhält man einen durchschnittlichen Differenzwert von 13.5 ($SD = 6.7$), was auf die 20 Fälle hochgerechnet in etwa einem Verhältnis der beiden Stoppregeln von 17:3 entspricht. Im Börsenszenario ergibt sich ein Differenzwert von 12.0 ($SD = 6.5$), also ein Verhältnis der beiden Stoppregeln von 16:4. Dieser Unterschied stellt eine Tendenz dar, ist jedoch im t-Test nicht signifikant ($p = 0.17$).

Entscheidungsverhalten

Die Basisrate für erfolgreiches Abfüllen / Investieren beträgt ein Drittel. Da die beiden Alternativen X und Y in ihren Ausprägungen nicht voneinander abhängig sind, ergibt sich somit ein Erwartungswert von 5/9 (56%), in denen wenigstens eine der beiden Alternativen positiv ist. Die Anzahl der positiven Entscheidungen der Versuchspersonen trifft in beiden Experimentalbedingungen in etwa diesen Wert: Im Trinkwasserszenario wird im Mittel in 12.6 der 20 ersten Durchgänge das Abfüllen gewählt (63%), im Börsenszenario im Mittel in 12.1 Durchgängen (61%). Der Unterschied zwischen beiden Varianten ist nicht signifikant ($p = .14$), die Ressource Zeit evoziert also eine allerdings nur geringe Tendenz zu häufigerem Investitionsverhalten als die Ressource Geld.

Deutlich stärker wird der Unterschied, wenn man diejenigen Fälle betrachtet, in denen sich eine Versuchsperson für genau eine der beiden Alternativen mit konfligierender Information konfrontiert sah (und für die andere Alternative keine positive Information hatte, so dass entweder nur die Alternative mit konfligierender Information in Frage kam oder gar keine). In diesen Fällen ist eine Entscheidung für die „wackelige“ Alternative substantiell riskant. Im Börsenszenario fanden sich 89 solcher Fälle, und die Versuchspersonen entschieden sich lediglich in 27% dieser Fälle dafür, ihr Geld trotz der substantiellen Unsicherheit in den entsprechenden Optionsschein zu investieren. Im Trinkwasserszenario kam eine Situation mit konfligierender Information 143mal vor. Hier entschieden sich die Versuchspersonen wesentlich häufiger, nämlich in 54% der Fälle dafür, die Wasserladung abzufüllen. Bei der Ressource Zeit verhalten sie sich somit wesentlich riskanter als bei der Ressource Geld.

Diskussion

Der grundsätzliche Befund dieses explorativen Experiments besteht darin, dass die Versuchspersonen im Wesentlichen bei der Ressource Zeit ein ähnliches Verhaltensmuster im Umgang mit Information und im Risikograd der Entscheidungen zeigen wie bei der Ressource Geld.

Allerdings setzen die Versuchspersonen im Trinkwasserszenario etwa ein Viertel mehr an Tests ein, und zwar bei den Cues, die nicht die validesten sind. Auf diese Weise erhöht sich – ohne dass der allgemeine Primat des *One Reason Decision Making* gebrochen wurde – die Anwendung einer More-Reason Stopping Rule. Bei acht der 60 Personen führt das dazu, dass die Majoritätsstrategie von der ORStR (Geld) zur MRStR (Zeit) kippt. Da bei 70% der Probanden die Majoritätsregel jedoch stabil bleibt, interpretieren wir dies für die gesamte Stichprobe eher im Sinne einer Tendenz, bei der Ressource Zeit eher schon einmal einen zusätzlichen Test durchzuführen als bei der Ressource Geld in einen zusätzlichen Indikator zu investieren. (Die individuell präferierte Stoppregele dominiert klar den Effekt der Bedingung.)

Ähnlich sieht das Bild bei der Entscheidungsregel aus: Hier besteht zunächst die (statistisch allerdings nicht signifikante) Tendenz, im Trinkwasserszenario etwas häufiger das Abfüllen zu wählen als im Börsenszenario das Geld auf einen Optionsschein zu setzen. Die sich daraus andeutende Interpretation, dass die Versuchspersonen mit der Ressource Zeit etwas grosszügiger umgehen als mit der Ressource Geld, wird substantiell durch den Befund unterstützt, dass sie im Falle konfligierender Information im Trinkwasserszenario deutlich häufiger bereit sind, eine Wasserladung abzufüllen, als dass sie im Börsenszenario investieren. Dies könnte

ein allgemeiner Unterschied im Umgang mit den beiden Ressourcen sein, auf den zu achten es sich in künftigen Experimenten lohnt.

Es gibt jedoch – und damit kommen wir zu den Limiten dieses explorativen Experiments – eine denkbare Alternativinterpretation für die gefundenen Unterschiede, und die liegt in der fixen Reihenfolge der beiden Experimentalbedingungen: Aus Gründen einer konsistent einheitlichen Instruktion führten alle Versuchspersonen zuerst das Trinkwasserszenario durch, bevor sie das Börsenszenario erarbeiteten. Hier könnte sich im Hinblick auf den Umfang der Informationssuche eine Art Lern- oder Erfahrungseffekt eingestellt haben. Es würde sich deswegen sicher anbieten, das Experiment sicherheitshalber noch einmal in anderer Reihenfolge der Bedingungen zu wiederholen. Wenn auch nicht auszuschliessen, so erscheint uns diese Alternativinterpretation jedoch sehr unwahrscheinlich, da sich ja gleichzeitig zeigte, dass die Versuchspersonen in der Tendenz im zuerst durchgeführten Szenario die klarere Stoppregel aufwiesen. Bei einem Lerneffekt sollte das aber genau anders herum sein, die Strategie sollte sich eher im zweiten Teil des Experiments ergeben.

Als erstes Experiment, welches die wichtige Alltagsressource Zeit in die Forschungsthematik „Informationssuche beim Entscheiden“ einbringt, kann diese Untersuchung nicht mit endgültigen und definitiven Befunden aufwarten. Es zeigt sich jedoch, dass diese Ressourcen zumindest in den Grundzügen miteinander vergleichbar sind. Ob sich die gefundenen Unterschiede hinsichtlich Grosszügigkeit / Sparsamkeit im Umgang mit den Ressourcen wirklich erhärten lassen, muss zukünftige Forschung zeigen.

Uns ist es dabei wichtig, der unterschiedlichen Charakteristik von Zeit und Geld im Hinblick auf Kumulierbarkeit und Einsatz dieser Ressourcen Rechnung zu tragen: Wie in der Einleitung ausgeführt, ist Zeit eine Ressource, die kontinuierlich vergeht und selten als unmittelbarer Nutzen im Fokus der Entscheidung steht (jeweils im Unterschied zu Geld). Das gilt es auch im Experiment zu berücksichtigen, und uns scheint dies mit dem Trinkwasserszenario im Grundzug gelungen zu sein. Allerdings ergeben sich dadurch – betrachtet aus dem rein numerischen Blickwinkel – etwas unterschiedliche Kosten-Nutzen-Erwartungen bei den Entscheidungen. Diese ermöglichen zusätzliche Alternativinterpretationen (vor allem solange die Befundlage zum Thema „Ressource Zeit“ noch derart dünn ist). Sie auszuräumen, würde ggf. eine leichte Anpassung des Börsenszenarios bedeuten, z.B. indem man einen konstanten Betrag von 10'000 Franken pro Durchgang in einen Optionsschein investiert und die Informationskosten erst im Nachhinein abgezogen werden.

Es gibt jedoch auch subtilere Unterschiede in den beiden Ressourcen, die fundamentalere Natur sind und sich durch Anpassungen in den Szenarien nicht ausräumen lassen: Beispielsweise erscheint es uns plausibel anzunehmen, dass das Erleben von Misserfolg bei beiden Ressourcen unterschiedlich sein könnte: Beide Szenarien-Typen charakterisieren sich durch unterschiedliche Kostenstrukturen bezüglich der "False Alarms" (Wahl von X oder Y obwohl negativ) und der "Missed Hits" (Wahl der 3. Alternative, wobei X oder Y oder beide positiv gewesen wären). Ein Verlust bedeutet im Trinkwasserszenario bei einem False Alarm 10 Stunden falsch investierte Zeit, bei einem Missed Hit verliert man nichts (ausser der eingesetzten Zeit für das Wassertesten). Verluste im Börsenszenario bedeuten bei einem False Alarm ein Verlieren der eingesetzten Betrags (bis zu CHF 10'000.-), bei einem Missed Hit aber das Verlieren des Doppelten des einsetzbaren Betrags (= 2/3 des möglichen Gewinns, da

lediglich ein Drittel auf das Gewinnkonto "gerettet" wird). Ein Missed Hit ist also eigentlich im Börsenszenario schwerwiegender. Das sollte dann eigentlich entweder zu einer intensiveren Informationssuche oder zu risikoreicherem Verhalten im Börsenszenario führen (um den Missed Hit zu vermeiden), beides ist jedoch nicht der Fall. Offenbar wird die Struktur des Misserfolgs nicht in dieser Weise von den Versuchspersonen wahrgenommen. Vielmehr verhalten sich die Versuchspersonen im Börsenszenario im Zweifelsfall (konfligierende Information) weniger riskant, und sie investieren auch weniger in Cue-Informationen. Beides würde sich erklären, wenn man annimmt, dass sie das Transferieren des Betrags auf das Gewinnkonto als Gewinn erleben (und nicht als möglicherweise entgangenen Gewinn und somit als Verlust). Im Trinkwasserszenario kann das Wegschütten beider Ladungen aber schwerlich als ein Gewinn erlebt werden, sondern maximal als das kostenlose Vermeiden eines substantiellen Zeitverlustes.

Diese strukturellen Unterschiede zwischen der Ressource Zeit und der Ressource Geld scheinen uns in der Sache selbst zu liegen und sind durch Parallelisierung im Experiment nicht oder nur unzureichend aufzuheben. Deshalb mag es nicht nur in diesem Experiment, sondern generell schwierig sein zu beurteilen, ob die beobachteten Verhaltensunterschiede allein auf die unterschiedlichen Ressourcen in den beiden Szenarien oder auch auf Unterschiede zwischen den Kosten-Nutzen-Strukturen zurückzuführen sind. Die Beantwortung dieser Fragen steht noch aus, und es wird weiter verfeinerte experimentelle Szenarien brauchen, um zu klaren Erkenntnissen über den Umgang mit der Ressource Zeit im Entscheidungsverhalten zu kommen.

Literatur

- Christen, S., Zurbriggen, S., Hausmann, D., & Läge, D. (2006). Trinkwasser- und Börsenszenario: Aufbau, Ablauf und technische Spezifikation einer Experimentalumgebung zur entscheidungsvorbereitenden Informationssuche in zeitlichen Verläufen. AKZ-Forschungsbericht 43, Zürich: Angewandte Kognitionspsychologie.
- Gigerenzer, G., & Goldstein, D. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650-669.
- Hausmann, D., & Läge, D. (2005). How the desired level of confidence influences the stopping rule of information search. In K. Opwis & I. Penner (Eds.), *Proceedings of KogWis05. The German Cognitive Science Conference 2005* (pp. 71-76). Basel: Schwabe.
- Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, 16, 366-387.
- Zurbriggen, S., Christen, S., Hausmann, D., & Läge, D. (2007). Informationssuche beim Entscheiden in zeitlich dynamischen Umwelten. AKZ-Forschungsbericht 44, Zürich: Angewandte Kognitionspsychologie.

Kapitel 6: Adaptivität der Informationssuche in statischen und in dynamischen Umwelten

Einleitung und Fragestellung

Mit der Entdeckung der Evolution als treibender Kraft für die kognitive Grundausstattung des Menschen ist in weiten Kreisen der Sozialwissenschaften die Überzeugung entstanden, dass soziale und kognitive Fähigkeiten stets im Kontext ihrer stammesgeschichtlichen Entwicklung zu modellieren seien. Die Palette von inzwischen als klassisch zu bezeichnenden fachwissenschaftlichen Entwürfen reicht dabei von den Sprachwissenschaften (z.B. Pinkers "The Language Instinct", 1994/1996) über die Psychologie (etwa Cosmides & Toobys „The Adapted Mind“, 1992) bis hinein in die Informatik (z.B. „Understanding Intelligence“, Pfeifer & Scheier, 1999). Die Folgerung, Fähigkeiten als adaptive Anpassungen an eine bestimmte Umwelt zu sehen (und sie entsprechend auch gemeinsam mit dieser Umwelt zu modellieren und in dieser experimentell zu untersuchen), findet in der Kognitionspsychologie ihren klarsten Ausdruck in der Konzeption einer „Adaptiven Werkzeugkiste“ (Gigerenzer & Todd, 1999). Mit dieser Toolbox wurde speziell für die Entscheidungsforschung, aber auch für die Kognitionspsychologie im Allgemeinen, eine Alternative zum klassischen Rahmenmodell eines weitgehend inhaltsunspezifisch nach generellen Verarbeitungsregeln operierenden kognitiven Systems konzipiert.

Anhand dieser Toolbox und dem damit verbundenen Einbezug der Umwelt konnte dann auch mit mathematischen Simulationen gezeigt werden, dass bestimmte Heuristiken in ganz bestimmten Klassen statistischer Umwelten zu erfolgreichen Entscheidungen / erfolgreichem Verhalten führen, wie z.B. die Rekognitionsheuristik (z.B. von Goldstein und Gigerenzer, 2002) oder die *Take The Best* Heuristik (Gigerenzer und Goldstein, 1996; 1999). In dem von Gigerenzer und Kollegen verwendeten Paradigma werden die statistischen Umwelten durch ein Set an Objekten definiert, welche auf einer Kriteriumsvariable einen Wert haben (z.B. Städtegrösse, Anzahl Autounfälle oder Attraktivitätswert von Personen) und ein Cue Profil aufweisen (Czerlinski, Gigerenzer, & Goldstein, 1999). Czerlinski et al. untersuchten so 20 verschiedene Umwelten, welche zwischen 11 und 395 Objekten und mit 3 bis 18 Cues variierten (siehe auch neuere Studie von Hogarth & Karelaia, 2006). Zahlreiche experimentelle Studien zeigen auch, dass Personen diese Heuristiken in den bestimmten Umwelten meistens anwenden, v.a. unter Zeitdruck (Rieskamp & Hoffrage, 1999) oder wenn die Informationen etwas kosten (Bröder, 2000b; Bröder & Schiffer, 2003a, 2003b; Newell & Shanks, 2003; Rieskamp & Hoffrage, 1999), aber eben nur meistens und nicht immer (Bröder, 2000a; Newell, Weston, & Shanks, 2003). Relativ wenig Wissen existiert bislang darüber, wie Personen die einzusetzenden Strategien lernen (siehe dazu Rieskamp & Otto, 2006; Rieskamp 2008) und wie sie diese Strategien in Abhängigkeit zur jeweiligen Umwelt auswählen (z.B. in Bezug auf Änderungen in den Korrelationen der Cues, Fasolo, Misuraca, & McClelland, 2003).

Interessant ist vor diesem Hintergrund vor allem die Wahl von solchen frugalen Urteilsheuristiken, die, um eine Entscheidung zu treffen, die potenziell vorhandenen Informationen nicht vollständig ausschöpfen, sondern mit partikulärer Information auskommen. Solche Heuristiken machen – wollen sie effizient sein – beträchtliche Annahmen hinsichtlich der statistischen Struktur der jeweiligen Umwelt: Nur innerhalb des richtigen statistischen Settings erweisen

sich frugale Heuristiken nämlich als leistungsfähig, und zwar dadurch, dass sie die statistischen Abhängigkeiten von Informationen ausnutzen und damit mit einem Teil derselben bereits zu einer hinlänglich guten Entscheidung führen können. Hat man hingegen keine Annahmen über die statistische Struktur der jeweiligen Umwelt, so ist es günstiger, die potentiell vorliegenden Informationen reichhaltiger auszuschöpfen (verbunden mit kognitiv aufwändigeren Verarbeitungsregeln), um Fehler aufgrund des Ignorierens von Informationen zu vermeiden.

Um sich für den Einsatz einer Urteilsheuristik zu entscheiden, bräuchte es also a) eine gewisse Einsicht in die statistische Struktur der Umwelt und b) ein Grundverständnis hinsichtlich der Wirkweise der Heuristiken. Mit diesen Requisiten im Gepäck lässt sich dann abschätzen, ob eine (die Performanz vermutlich etwas reduzierende) frugale Heuristik in einer bestimmten Situation angemessen ist oder ob es sich lohnt, eine umständlichere Verarbeitung von reichhaltigerer Information vorzunehmen. Doch basiert die Wahl der Strategie wirklich auf derartigen Metakognitionen? Rieskamp (2006) kommt in diesem Zusammenhang aufgrund seiner empirischen Arbeiten zu der interessanten Schlussfolgerung, dass Menschen Strategien nicht etwa durch einen antizipierten Trade-off von Geschwindigkeit und Genauigkeit treffen, sondern eine Strategie im Prozess selbst zu wählen lernen, indem sie nach und nach Einsicht in die statistische Struktur der aktuellen Umwelt gewinnen. Er zeigte dies anhand zweier unterschiedlicher Umwelten, einer mit kompensatorischen Gewichten probabilistischer Hinweisreize und einer mit non-kompensatorischen. In der kompensatorischen Umwelt konnten also die weniger guten Cues den validesten noch „überstimmen“ (weshalb es sich ggf. lohnte, sich mit ihnen zu beschäftigen), während das in der non-kompensatorischen Umwelt nicht der Fall war: Hier lohnte es, die Entscheidung immer aufgrund des validesten Cues zu treffen. Rieskamps Befunde zeigen, dass seine Versuchspersonen anfänglich eine Präferenz für eine kompensatorische Strategie haben, d.h. dass sie mit Vorliebe die Informationen integrieren. Sie scheinen jedoch im Experiment die statistische Struktur zu lernen und stellen dann ggf. ihre Strategie um. Dabei fand er allerdings einen „inertia effect“, d.h. Personen bleiben zu lange an der zuvor erfolgreich angewendeten Suchstrategie haften, auch wenn die Umwelt sich verändert (von kompensatorisch zu non-kompensatorisch und vice versa) und eine andere Strategie erfolgreicher wäre.

Dieser Befund führt zu der Fragestellung, wie adaptiv denn die Wahl einer Strategie ist. Aus Rieskamps Daten lässt sich dies nur sehr bedingt beantworten, denn – gemäß den Zielen seiner Untersuchung – war die Art der statistischen Umwelt ja erst innerhalb des Experiments zu entschlüsseln. Anders könnte es aussehen, wenn den Personen von vornherein klar ist, in was für einer Umwelt sie sich bewegen.

Payne & Bettman (2000) stellen als Zusammenschau der bis dahin vorliegenden Literatur fest, dass sich Personen im wesentlichen bei der Wahl ihrer Strategien adaptiv hinsichtlich der Umwelt verhielten, dass man jedoch auch Fehler konstatieren könne, die darin bestünden, dass Personen auf die falsche Heuristik zugriffen. Diese recht generelle Einschätzung ist, sofern es um Entscheidungen und nicht um Wahlpräferenzen geht, allerdings nur auf Experimente gestützt, in denen unterschiedlichste Parameter der statistischen Umwelten variiert wurden (wie z.B. die Informationssuchkosten), jedoch nicht die statistische Umwelt selbst. Insofern fehlen hier also die zur Beantwortung der Frage notwendigen Evidenzen.

Adaptives Verhalten bei unterschiedlichen statistischen Umwelten

Im Folgenden werden wir ein Experiment präsentieren, in welchem – innerhalb desselben inhaltlichen Szenarios – zwei völlig unterschiedliche statistische Umwelten vorkommen. Beide Umwelten sind als nonkompensatorische Umwelten gestaltet und die Cues sind unkorreliert. Ihr statistischer Unterschied zeigt sich in einem anderen Aspekt, und zwar der logischen Verknüpftheit von Ereignissen. Die erste Umwelt entspricht derjenigen, welche in der experimentellen Entscheidungsforschung bis anhin die Regel war, z.B. im Börsenszenario (Bröder, 2000b; Bröder & Schiffer, 2006; Läge & Hausmann, 2007; Newell, Rakow, Weston, & Shanks, 2004; Newell & Shanks, 2003; Newell et al., 2003; Rakow, Newell, Fayers, & Hersby, 2005). In dieser Umwelt sind die Informationen, auf deren Basis jeweils eine Entscheidung zu treffen ist, nicht über die Versuchsdurchgänge hinweg miteinander verbunden und die Entscheidung ist somit ein statisches Einzelereignis. In einer zweiten Umwelt jedoch sind die Informationen logisch verbunden, d.h. einmal vorliegende Daten lassen sich auch für spätere Entscheidungen noch heranziehen. Diese Umwelt ist somit zeitlich dynamisch. Da Bröder fand, dass Personen sich adaptiver verhalten, wenn die semantische Einkleidung und optische Gestaltung des Bildschirms eine Umweltänderung andeutet (Bröder, 2005), wird in der zweiten Umwelt diese Zeitdimension sowohl inhaltlich eingeführt wie auch grafisch dargestellt. Auf dieser Zeitdimension sind die einzelnen Durchgänge platziert: Die Versuchspersonen werden gebeten, mit Hilfe von Tests die Qualität von Wasserlieferungen an eine Trinkwasser-Abfüllstelle zu beurteilen. Die beiden parallel verfügbaren Wasserlieferungen stammen jeweils aus denselben Quellen, von denen bekannt ist, dass sie zeitweise gute Qualität besitzen, zeitweise sich jedoch nicht für Trinkwasser eignen. Die Versuchspersonen können somit – natürlich mit Abstrichen an die Validität – Testresultate aus vorhergehenden Durchgängen mit für ihre Evaluation heranziehen. Die aufgedeckten Cue-Informationen der letzten vier Runden bleiben auf dem Information Board sichtbar, wodurch mögliche Gedächtnisbelastungen ausgeschlossen werden, sofern Personen vergangene Informationen in die aktuelle Entscheidung mit einbeziehen möchten.

Die kritische abhängige Variable, die sich in diesen beiden statistischen Umwelten unterscheiden sollte, ist der Umgang mit Information: Im Szenario der logisch verbundenen Ereignisse stehen auch die Testresultate aus der Vergangenheit mit gewisser „Restvalidität“ zur Verfügung. Die Versuchspersonen können hier also sparsamer mit der Durchführung von Tests umgehen als in der Umwelt logisch unverbundener Ereignisse. Um zu verhindern, dass einfach sämtliche Tests durchgeführt werden (pro Wasserlieferung stehen jeweils bis zu vier Tests unterschiedlicher Validität zur Verfügung), wird die Durchführung mit einem Kostenaufwand verknüpft: Die Durchführung eines Tests für eine der beiden parallel verfügbaren Wasserlieferungen dauert eine „virtuelle“ Arbeitsstunde. Entscheidet sich die Versuchsperson dafür, eine Lieferung in Wasserflaschen abzufüllen, so dauert dies zehn Stunden. (Als Alternative kann sie beide verfügbaren Wasserlieferungen wegschütten lassen, was sie keinen Zeitaufwand kostet.) Die Versuchsperson hat für das Experiment eine fixe Arbeitszeit (von 240 Stunden) zur Verfügung, so dass die Durchführung vieler Tests die Zahl der Möglichkeiten begrenzt, Trinkwasser abzufüllen. Da am Ende des Experiments die Anzahl erfolgreich abgefüllter Trinkwasserlieferungen zählt, ist somit das Durchführen vieler Tests mit einem

erheblichen Ressourcenaufwand verbunden. Mit dieser Ressource, so nehmen wir an, werden die Versuchspersonen so sparsam umgehen wie ihnen geraten erscheint.

Zentral für die Anpassung an die jeweilige statistische Umwelt ist also die Stopppregel, die die Versuchspersonen beim Durchführen von Tests verwenden: In Frage kommt das Stoppen nach einem einzigen durchgeführten Test (was wir als One Reason Stopping Rule, ORStR) bezeichnen wollen, das Weitertesten trotz vorliegender Information (More Reason Stopping Rule, MRStR) sowie das Abfüllen oder Wegschütten einer Wasserlieferung, ohne dafür einen Test durchgeführt zu haben (No New Reason Stopping Rule, NNRStR). Da im Experiment die Basisrate für trinkbare Wasserlieferungen auf ein Drittel gesetzt ist (was der Versuchsperson in der Einleitung mitgeteilt wird), ist die NNRStR eigentlich nur zu erwarten, wenn auf Testinformationen aus der Vergangenheit vertraut wird, also in der Umwelt mit logisch verbundenen Ereignissen. Deshalb wird das „New“ in NNRStR eingefügt, weil ja durchaus Testergebnisse aus der Vergangenheit herangezogen werden können, aber dabei keine neue Information gekauft wird.

Das Szenario – welches in zwei Forschungsberichten (Christen, Zurbriggen, Hausmann, & Läge, 2007; Zurbriggen, Hausmann, Christen, & Läge, 2007) ausführlichst beschrieben wird – arbeitete mit der kritischen Ressource Zeit (statt mit virtuellen Geldbeträgen, wie es in der Entscheidungspsychologie der Regelfall ist). Ein ansonsten weitgehend parallel aufgebautes, aber mit der Ressource Geld operierendes Börsenexperiment (Zurbriggen, Christen, Hausmann, & Läge, 2008) hat gezeigt, dass Personen unterschiedlich auf die zwei statistischen Umwelten reagieren: Erstens suchen sie mehr Informationen in einer Umwelt, die lediglich aus Einzelereignissen besteht, als in einer Umwelt, die aus Serien von Ereignissen besteht. Und zweitens entscheiden sie sich häufiger ohne (neue) Information, wenn die Ereignisse zeitlich miteinander verknüpft sind und Personen die Entscheidungen nicht als eine Reihe von Einzeldurchgängen erleben, sondern in einen Zeitfluss eingebunden. Diese Befunde basieren auf Unterschieden von einer Versuchspersonengruppe, welche immer mit derselben statistischen Umwelt ins Experiment startete. In dem im Folgenden berichteten Experiment im Trinkwasserszenario soll nun einerseits untersucht werden, wie adaptiv zwei Versuchspersonengruppen auf die beiden statistischen Umwelten reagieren und andererseits wie adaptiv die Personen bei einem Wechsel von einer statistischen Umwelt zur anderen, sich verhalten – und zwar in beide Richtungen.

Der Wechsel zwischen den beiden Umwelten wird den Personen durch die Versuchsinstruktion am Beginn des nächsten Durchgangs bekanntgegeben (und dabei in einen inhaltlichen Kontext gesetzt), so dass die Personen auch die Gelegenheit haben, ihre Strategie unmittelbar an die neuen Verhältnisse anzupassen. Es gibt kein direktes Feedback über die Korrektheit einer gerade getroffenen Entscheidung, sondern die Versuchspersonen werden alle drei bis fünf Runden über die Zahl der erzielten Treffer informiert, so dass während der Durchgänge kein erfolgsabhängiges Lernen einer Strategie möglich ist.

Es stellen sich hier folgende, zunächst explorativ formulierte Forschungsfragen: Wie adaptiv verhalten sich Personen in Bezug auf unterschiedliche statistische Umwelten? Als abhängige Variable sollen die Stopppregel und die Entscheidung der Personen im Experiment ausgewertet werden. In Bezug auf die Stopppregel nehmen wir an, dass die Personen diese an die beiden statistischen Umwelten anpassen. In der Umwelt mit den verbundenen Ereignissen sollte die

vergangene Information immer noch einen Wert haben, da die Information von „gestern“ etwas über den Zustand von heute aussagt. Daher formulieren wir 1.) die Hypothese, dass die Verknüpfung von Ereignissen dazu führt, dass insgesamt weniger Informationen gesucht werden. Dies bedeutet, dass 2.) die Informationssuche bei verbundenen Ereignissen früher gestoppt werden müsste als bei unverbundenen Ereignissen: Die Personen sollten einerseits in der Umwelt der verbundenen Ereignisse häufiger ohne aktuelle Information entscheiden und andererseits häufiger mit einer Information zufrieden sein (weil ihnen ja in beiden Fällen noch mehrere Informationen aus der Vergangenheit zur Verfügung stehen) als in der Umwelt der unverbundenen Ereignissen. Wir erwarten zudem 3.) ein adaptives Entscheidungsverhalten: Bei den unverbundenen Ereignissen sollte die Wahl einer Alternative im Vergleich zum Vordurchgang eher zufällig verteilt sein. Bei den statistisch verbundenen Ereignissen hingegen sollte die Wahl von der vorherigen Wahl abhängig sein.

Als weiteres wird dann untersucht, wie die Personen auf die Veränderung der statistischen Umwelt reagieren. Dies geschieht, indem eine Gruppe von Versuchspersonen zunächst das erste Trinkwasserszenario mit einzelnen Wasserlieferungen aus jeweils verschiedenen, unbekannten Wasserquellen durchführt (T1P1 für „Trinkwasserszenario 1 in Position 1“) und danach mit dem zweiten Trinkwasserszenario konfrontiert wird, in dem die Wasserlieferungen jeweils aus denselben beiden Wasserquellen stammen (T2P2 für „Trinkwasserszenario 2 in Position 2“). Eine zweite Gruppe von Versuchspersonen beginnt mit Trinkwasserszenario 2 (T2P1) und wechselt danach in das Trinkwasserszenario 1 (T1P2). Aus diesem Design ergibt sich folgende Operationalisierung der oben gelisteten Fragestellungen:

1. Adaptives Verhalten bei unterschiedlichen statistischen Umwelten (between subjects)

Welchen Einfluss hat die statistische Umwelt (unverbundene Ereignisse T1 vs. verbundene Ereignisse T2) auf das Stoppverhalten und die Entscheidung? Um diese Frage zu untersuchen, werden die beiden Gruppen untersucht, wobei die eine Gruppe das Experiment mit der Umwelt der unverbundenen Ereignisse beginnt, die andere mit den verbundenen. Da die beiden statistischen Umwelten somit am Anfang des Experimentes stehen, kann das Verhalten in den beiden Umwelten direkt verglichen werden.

Aufgrund der in T2 verfügbaren Informationen aus der Vergangenheit ist zu erwarten, dass die Versuchspersonen in dieser Bedingung deutlich weniger Tests durchführen. Entweder die Personen reduzieren ihre Informationssuche in T2, indem sie sich häufig ohne aktuelle Information entscheiden oder sie begnügen sich häufiger mit einer statt mehreren Informationen. Diese Möglichkeit, Testresultate aus der Vergangenheit heranzuziehen, sollte daher in der Summe zu mehr NNRStR und ORStR's (auf Kosten von der MRStR) führen. Beim Entscheidungsverhalten in der Umwelt der statistisch verbundenen Ereignissen sollte die Wahl einer bestimmten Alternative die Wahrscheinlichkeit erhöhen, dieselbe Wahl in der nächsten Runde zu treffen: Die Wahlwiederholung der beiden Alternativen X und Y ist in dieser Umwelt höher als der Umwelt der Einzelereignisse.

2. Adaptation beim Wechsel der statistischen Umwelt (within subjects)

Wenn die Versuchspersonen von einer statistischen Umwelt in die andere wechseln, dann sollten sie dieselben, oben ausgeführten Unterschiede im Verhalten zeigen wie Personen, die jeweils eine dieser Umwelten bearbeiten. Um das zu überprüfen, absolvieren die beiden oben

genannten Gruppen als zweites nun die Umwelt, mit welcher sie noch nicht konfrontiert wurden: Wer mit T1P1 ins Experiment gestartet ist, erlebt einen Wechsel von einer Umwelt, welche zuerst aus unverbundenen Ereignissen bestand, zu einer aus verbundenen Ereignissen (T1P1 - T2P2). Und diejenigen Versuchspersonen, die zuerst mit verbundenen Ereignissen konfrontiert wurden, erleben einen Wechsel zu unverbundenen Ereignissen (T2P1 - T1P2).

Methode

Stichprobe

Es nahmen 60 Versuchspersonen teil (30 Männer und 30 Frauen), darunter 38 Psychologie-Studierende der Universität Zürich. Das Durchschnittsalter betrug 30.3 Jahre (Range 19.8-60.8, SD = 8.9). Zwei Stichproben führten das Experiment unterschiedlich durch, wobei die Grösse jeder Stichprobe $N=30$ war. Die beiden Versuchspersonengruppen unterschieden sich weder im Alter ($t=-.81$, $p=.42$), noch im Geschlecht (Chi-Quadrat = .27, $p=.61$) oder der Ausbildung¹ (Chi-Quadrat = 2.8, $p=.09$) voneinander.

Die Psychologie-Studierenden müssen im Rahmen ihrer Ausbildung an der Universität Zürich mindestens fünf Stunden an Experimenten teilnehmen, an welchen ist ihnen dabei frei gestellt. Die Nicht-Studierenden nahmen ohne Gutschrift von Versuchspersonenstunden oder einer anderen Form der Entschädigung teil. Um die Versuchspersonen zu einer überdurchschnittlichen Leistung zu motivieren, wurden unter dem besten Drittel aller Versuchspersonen zwei Gewinne von je CHF 100.- verlost. Die Experimentalsoftware war in Microsoft Visual Basic 6.0 programmiert und das Experiment wurde vollständig auf einem PC mit Windows XP-Betriebssystem durchgeführt.

Vorgehen

Wie weiter oben bereits erläutert, entspricht das Konzept des Experiments einem 2x2 Design, wobei die Position der statistischen Umwelten zwischen den Personen variiert wird. So kann einerseits das adaptive Verhalten bei unterschiedlichen statistischen Umwelten gemessen werden (between subjects), andererseits aber auch die Adaptation an den Wechsel der Umwelt (within subjects). Als Szenario wurde das Trinkwasserszenario mit den beiden statistischen Umwelten T1 und T2 herangezogen (siehe Beschreibung oben und die Illustrationen im Experiment zu den beiden statistischen Umwelten in Abbildung 18 und 19). Die vier zur Verfügung stehenden Tests hatten die Validitäten .75, .80, .85 und .90 (wobei die Validitäten in ganzen Zahlen dargestellt wurden). Dabei bezog sich ein Test immer auf eine von zwei verfügbaren Wasserlieferungen. Das Durchführen eines Testes kostete eine virtuelle Arbeitsstunde, das Abfüllen einer Wasserladung dauerte jeweils zehn Stunden, das Wegschütten (im Szenario mit der Option „Keine Abfüllung“ bezeichnet) benötigte keine Zeit. Insgesamt standen der Versuchsperson 240 Stunden Spielzeit zur Verfügung, und sie sollte so viele Wasserlieferungen wie möglich in dieser Zeit abfüllen (wobei natürlich nur die „guten“ gezählt werden). Die Information Boards der beiden statistischen Umwelten sind in Abbildung 20 dargestellt.

¹ Die Ausbildung wurde nach zwei Gruppen kategorisiert: Berufslehre und Universitätsausbildung.

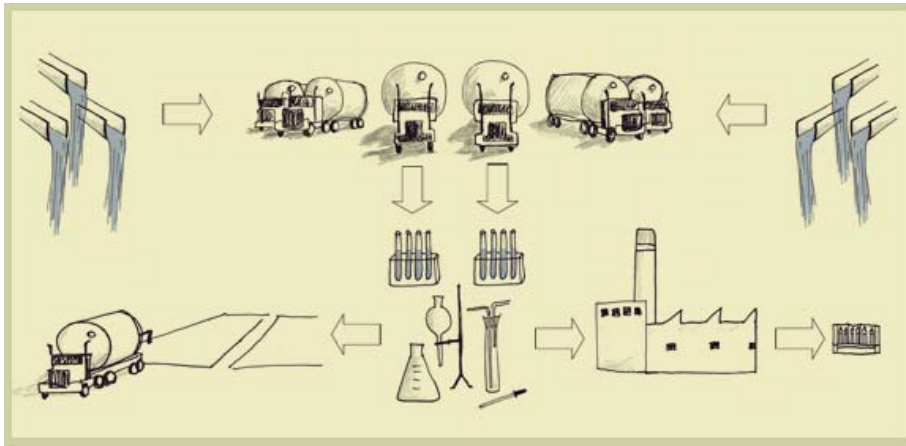


Abbildung 18: Illustration des Trinkwasser-Szenarios zur Einführung in statistische Umwelt der Einzelereignisse (T1).

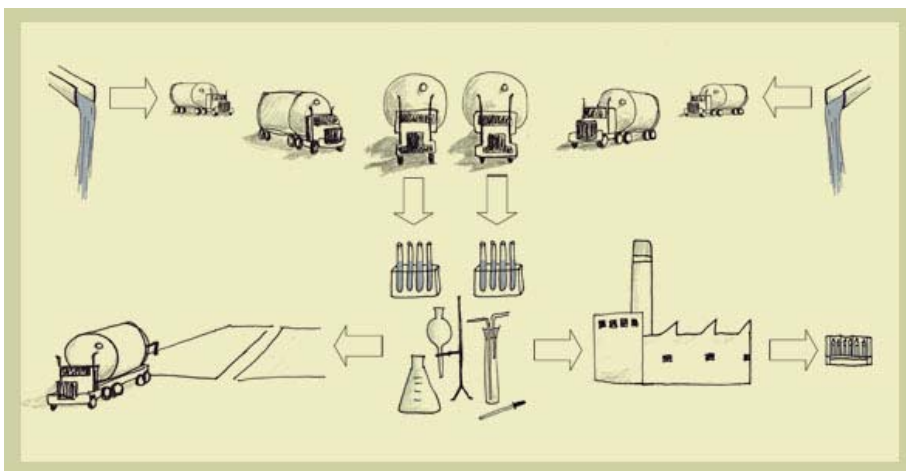


Abbildung 19: Illustration des Trinkwasser-Szenarios zur Einführung in die statistischen Umwelten der verbundenen Ereignisse (T2).

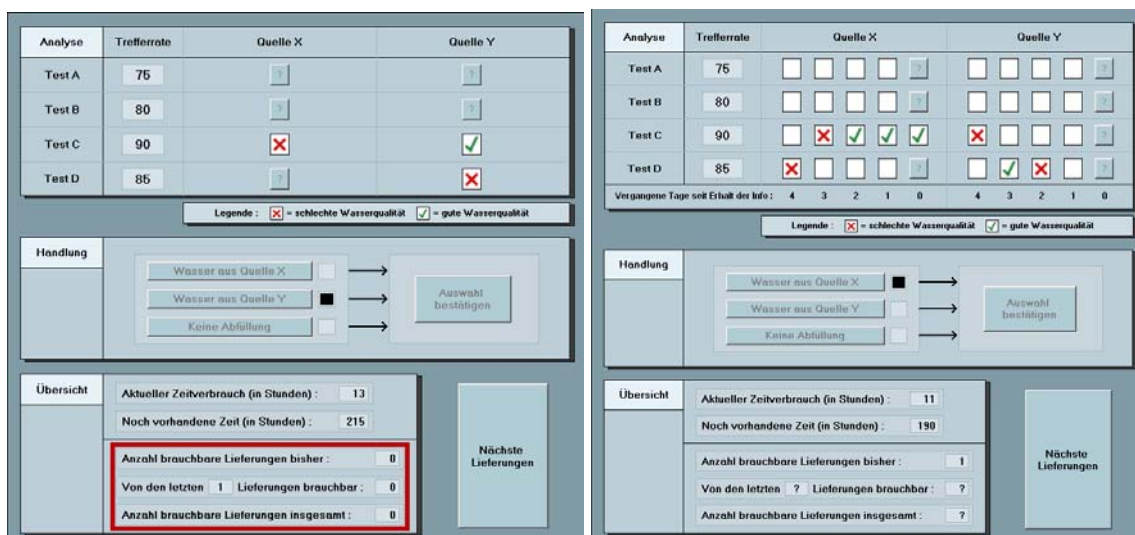


Abbildung 20: Information Boards der beiden Bedingungen (links: Einzelereignisse T1, rechts: mit Zeitdimension T2)

Resultate: Überblick

In einem ersten Schritt werden in einem Überblick die wichtigsten Befunde und deren Auswertung in Bezug auf das Stopp- und Entscheidungsverhalten zusammenfassend dargestellt. In einem zweiten Schritt werden die Befunde detailliert aufgelistet und diskutiert. Um die Wahrnehmung der Unterschiedlichkeit der beiden statistischen Umwelten und das Verhalten darauf zu analysieren, sind die Ergebnisse des Vergleichs zwischen T1 und T2 in der ersten Position im Experiment relevant (siehe Tabelle 31 Fragestellung F1). Und um die Adaptation bei einem Wechsel der Umwelt zu beantworten, sind folgende zwei Vergleiche von Interesse: der Wechsel von T1P1 zu T2P2 und von T2P1 zu T1P2 (siehe Tabelle 31 Fragestellung F2).

Tabelle 31: Fragestellung adaptives Verhalten an unterschiedlich statistische Umwelten (F1) und Fragestellung Adaptation bei Wechsel (F2), wobei T1 das Trinkwasserszenario der unverbundenen Einzelereignisse, T2 das Trinkwasserszenario der verbundenen Ereignisse, P1 die erste Position im Experiment und P2 die zweite Position im Experiment ist.

	T1	T2			T1	T2			T1	T2
P1	↔	↔			P1				P1	
	F1	F1				F2				F2
P2					P2				P2	
						F2				F2

Stoppverhalten

Ob die Versuchspersonen vergangene Informationen in ihre Informationssuche mit einbezogen oder nicht, lässt sich beim Stoppverhalten anhand der Anzahl durchgeführter Tests (siehe Tabelle 32) und anhand der Stoppregeln (siehe Tabelle 33) beschreiben.

Obwohl in der Umwelt der Einzelereignisse T1P1 eindeutig häufiger als in der Umwelt der verbundenen Ereignisse T2P1 Cues pro Durchgang gekauft wurden, wird diese Tendenz zwischen den beiden unabhängigen Stichproben nicht signifikant ($p=.09$). Dafür zeigt sich aber beim Wechsel von T1P1 zu T2P2 eine signifikante Veränderung der Cuekäufe ($p<.001$). Starten die Versuchspersonen das Experiment jedoch mit der Umwelt der verbundenen Ereignisse T2P1, dann verändert sich die Cuekaufhäufigkeit zu der zweiten Bedingung T1P2 hin nicht ($p=.29$).

Tabelle 32: Anzahl durchgeführter Tests pro Durchgang.

	T1	T2
P1	M = 2.8 SD = 1.2	M = 2.3 SD = 1.4
P2	M = 2.2 SD = 1.2	M = 2.4 SD = 1.2

In Bezug auf die Stoppregeln (siehe Tabelle 33) passen die Versuchspersonen die Informationssuche an die Umwelt der verbundenen Ereignisse an, indem sie häufig ohne neue Informationen (NNRStR) stoppen ($p<.001$). Dieser Befund findet sich auch bei einem Wechsel von T1P1 zu T2P2 ($p<.05$). Beim Wechsel von T2P1 zu T1P2 bleibt die Tendenz erhalten, aber das Resultat wird nicht signifikant ($p=.10$). Fasst man die NNRStR über die beiden Positionen zusammen (siehe Mittelwert in Tabelle 33), zeigt sich ein eindeutiger Einfluss der **statisti-**

schon Umwelt: Die Versuchspersonen entschieden sich in der Umwelt der verbundenen Ereignisse T2 häufiger ohne jegliche Information als in der Umwelt unverbundener Einzelereignisse T1 ($t=-3.0$; $p<.01$).

In den beiden statistischen Umwelten werden die beiden Stoppregeln ORStR und MRStR relativ stabil und somit unabhängig von der statistischen Umwelt oder derer Position von den Versuchspersonen angewendet. In beiden Umwelten wird die Informationssuche ein wenig häufiger nach einer Information gestoppt als nach mehreren. Aber wie die grossen Standardabweichungen erahnen lassen, sind diese Unterschiede auf Gruppenniveau nicht signifikant.

Tabelle 33: Angewendete Stoppregeln und die Summe der angewendeten NNRStR&ORStR in allen Durchgängen.

	T1			T2			T1	T2
Stopp	NNRStR	ORStR	MRStR	NNRStR	ORStR	MRStR	NNRStR & ORStR	NNRStR & ORStR
P1	M = 0.3 SD = 0.6	M = 13.8 SD = 12.4	M = 12.4 SD = 9.2	M = 3.3 SD = 5.7	M = 13.4 SD = 7.9	M = 10.5 SD = 6.8	M = 14.1 SD = 12.6	M = 16.7 SD = 10.1
P2	M = 1.8 SD = 4.2	M = 18.2 SD = 10.9	M = 9.7 SD = 10.0	M = 4.0 SD = 7.9	M = 14.5 SD = 11.5	M = 12.7 SD = 10.8	M = 20.0 SD = 10.8	M = 18.5 SD = 13.9
Mittelwert	1.1	16	11.1	3.6	14	11.6	17.1	17.6

Bei der Bearbeitung einer statistischen Umwelt unterscheidet sich die Summe der beiden Stoppstrategien ORStR und NNRStR nicht ($p=.38$). Bei beiden Wechseln ist die Summe der beiden Stoppstrategien in der zweiten Bedingung jeweils signifikant ($p<.001$) grösser als in der ersten Bedingung, was auf einen Positionseffekt hindeutet.

Entscheidungsverhalten

In den ersten 20 Runden zeigen die beiden Versuchspersonengruppen ein klares adaptives Verhalten an die beiden statistischen Umwelten in Bezug auf die Wahlwiederholungen (siehe Tabelle 34). In der Umwelt der statistisch verbundenen Ereignissen wird eine Entscheidung durch die vorherige Entscheidung stärker beeinflusst als in der Umwelt der Einzelereignisse ($p<.01$). Bei dem Wechsel von T1P1 zu T2P2 zeigt sich ein adaptives Verhalten in dieselbe Richtung ($p<.01$), bei dem Wechsel von T2P1 zu T1P2 finden wir die Adaption nur in der Tendenz ($p=.15$). Über die beiden Positionen gemittelt zeigt sich ein Einfluss der **statistischen Umwelt** auf die Wahlwiederholung ($t=-3.2$, $p<.01$).

Tabelle 34: Wahlwiederholung der Alternativen in den ersten 20 Runden.

	T1	T2
P1	M = 5.8 SD = 1.7	M = 7.3 SD = 3.0
P2	M = 6.5 SD = 3.6	M = 7.4 SD = 2.8
Mittelwert	6.2	7.3

Entscheiden sich die Versuchspersonen fürs Abfüllen von einer der beiden Wasserladungen (siehe Tabelle 35), dann präferieren sie in der Umwelt der unverbundenen Ereignisse T1P1 die Alternative X. In der Umwelt der verbundenen Ereignisse T2P1 hebt sich diese asymmetrische Präferenz jedoch wieder auf. Da sich diese Aufhebung der Asymmetrie in T2P2 aber nicht replizieren lässt und die Personen somit beim Wechsel von T1P1 zu T2P2 nicht adaptiv

auf die Umwelt reagieren, dafür aber beim Wechsel von T2P1 zu T1P2 die Asymmetrie wieder auftaucht, deutet dieses Resultat auf einen **Interaktionseffekt** von Position und statistischer Umwelt bei der Alternativenwahl hin. Die rechtspositionierte Alternative Y wurde nicht vernachlässigt, wenn die Ereignisse miteinander verbunden sind und als erstes bearbeitet wurden (T2P1). Diese Interaktion wird bei der Alternative Y signifikant ($F = 8.8$ $p < .01$, Partielles Eta-Quadrat .13).

Tabelle 35: Total Entscheidungen in den ersten 20 Runden.

Alternative	T1			T2			Mittelwert		
	X	Y	Keine Abfüllung	X	Y	Keine Abfüllung	X	Y	Keine Abfüllung
P1	M = 7.6 SD = 1.8	M = 5.0 SD = 2.0	M = 7.3 SD = 1.7	M = 6.3 SD = 2.3	M = 6.8 SD = 3.1	M = 6.6 SD = 2.8	7.0	5.9	6.9
P2	M = 7.5 SD = 2.7	M = 5.0 SD = 3.4	M = 7.5 SD = 2.9	M = 7.4 SD = 1.7	M = 4.2 SD = 2.5	M = 8.4 SD = 2.8	7.5	4.6	8.0
Mittelwert	7.6	5.0	7.4	6.9	5.5	7.5			

Detaillierte Resultate des Vergleichs (between subjects)

Adaptives Verhalten bei unterschiedlichen statistischen Umwelten T1P1 - T2P1

In dieser detaillierteren Auswertung ist zu überprüfen, inwiefern die beiden Versuchspersonengruppen unterschiedlich auf die beiden statistischen Umwelten reagierten. Dazu werden die beiden Szenarien T1P1 (Versuchspersonen, die mit den unverbundenen Ereignissen starten) und T2P1 (Versuchspersonen, die mit den verbundenen Ereignissen starten) miteinander verglichen. Bei den inferenzstatistischen Auswertungen wurden entsprechend t-Tests für unabhängige Stichproben mit zweiseitiger Hypothesenprüfung durchgeführt ($\alpha = .05$).

Stoppverhalten

Anzahl durchgeführter Tests

In der Umwelt der unverbundenen Einzelereignisse (T1P1) führte eine Person im Schnitt 69.2 Tests durch (SD=21.6), bei den verbundenen Ereignissen (T2P1) waren es hingegen nur 57.2 Tests (SD=23.8). Dieser in den Hypothesen erwartete Unterschied erweist sich im t-Test als signifikant ($t = 2.0$; $p < .05$).

Es wurden in beiden Umwelten im Mittel mehr Tests durchgeführt als für ein *One Reason Decision Making* nötig gewesen wäre: In T1P1 führte eine Person im Mittel 2.8 Tests pro Runde durch, und in T2P1 lag der Wert mit 2.3 Tests ebenfalls über der „Schallmauer“ von zwei Informationen (auch wenn der Test bei beiden Alternativen negativ ausfällt), ab der spätestens das More-Reason Decision Making beginnt (ohne signifikanten Unterschied; $t = 1.4$; $p = .18$). Schauen wir uns deswegen die Stoppregel etwas genauer an.

Stoppregel

Jeder Durchgang lässt sich einer der drei oben eingeführten Stoppregeln zuordnen, wenn man folgende definitorischen Präzisierungen hinsichtlich des eingesetzten Versuchsdesigns vornimmt: Wir klassifizieren als One Reason Stopping Rule (ORStR), wenn die Informationssuche nach dem ersten Test gestoppt wird oder wenn nach einem ersten negativen Test noch

genau ein weiterer Test durchgeführt wird, und zwar für die andere Alternative (da diese sich als positiv erweisen könnte). Entsprechend gilt als More Reason Stopping Rule (MRStR), wenn zu einer Alternative mehr als ein Test durchgeführt wird oder wenn nach einem ersten positiven Test auch noch die andere Wasserlieferung getestet wird. Bei der No New Reason Stopping Rule (NNRStR) wird ohne jegliche Zuhilfenahme von aktueller Information entschieden.

Untersucht man die Stoppregeln anhand dieser Klassifikationen, zeigt sich, dass in beiden statistischen Umwelten über alle Runden gesehen etwas häufiger anhand einer ORStR entschieden wurde (T1P1: $M=13.8$, $SD=12.4$; T2P1: $M=13.4$, $SD=7.9$) als anhand einer MRStR (T1P1: $M=12.4$, $SD=9.2$; T2P1: $M=10.5$, $SD=6.8$). Der Unterschied zwischen ORStR und MRStR ist jedoch in beiden Umwelten nicht signifikant (für T1P1 $t=0.35$; $p=.73$, für T2P1 $t=1.22$; $p=.23$). Das heisst, die Informationssuche wird in beiden Umwelten etwa gleich häufig nach einer wie nach mehreren Informationen gestoppt. Betrachtet man die individuellen Daten, dann findet man in beiden statistischen Umwelten etwa gleich viele One-Reason Decision Maker wie More-Reason Decision Maker. In T1P1 können 14 ORDM und 16 MRDM klassifiziert werden und bei T2P1 15 ORDM und 15 MRDM (siehe ausführliche Darstellung bei den intraindividuellen Auswertungen).

Es zeigt sich aber ein Unterschied zwischen den beiden statistischen Umwelten beim Stoppen ohne Informationssuche (NNRStR): In der Umwelt der Einzelereignisse wurde praktisch nie ohne Information gestoppt ($M=0.3$, $SD=0.6$) und wenn, dann wurde jeweils für eine Alternative X oder Y und nie für die Option „Keine Abfüllung“ entschieden. In der Umwelt der verbundenen Ereignisse wurde hingegen signifikant häufiger ($M=3.3$, $SD=5.7$, $t=-2.82$; $p<.001$) ohne Informationssuche entschieden und dabei auch mehrfach für „Keine Abfüllung“. Dieser Effekt ist allerdings, wie die grosse Standardabweichung bereits erahnen lässt, vornehmlich auf wenige Versuchspersonen zurückzuführen (vgl. erste Spalte in Tabelle 37). Beim Durchdenken der möglichen Strategien hätte man erwarten dürfen, dass noch mehr Versuchspersonen in T2P1 von dieser Option Gebrauch gemacht hätten, da im Falle negativer Testergebnisse bei der Wasserlieferungen aus ein- und derselben Quelle man die Wahrscheinlichkeit, wieder einen positiven Test zu bekommen, mit einigen weggeschütteten Lieferungen ja etwas erhöhen kann. Zumindest einige der Versuchspersonen haben die statistischen Abhängigkeiten in dieser Umwelt auf derartige Weise genutzt.

Wir nahmen in den Hypothesen an, dass beim Heranziehen der Testresultate aus der Vergangenheit die Summe von NNRStR und ORStR's in der Umwelt der statistisch verbundenen Ereignissen höher sein sollte als in der Umwelt der Einzelereignisse (und zwar auf Kosten der MRStR). Dies ist zwar der Fall (T2P1: $M=16.7$, $SD=10.1$ und T1P1: $M=14.1$, $SD=12.6$), aber nicht signifikant ($t=-.86$, $p=.38$).

Entscheidungsverhalten

Zur Auswertung des Entscheidungsverhaltens werden die ersten 20 Durchgänge einer jeden Person betrachtet, um trotz unterschiedlicher Anzahl gespielter Durchgänge sicherzustellen, dass je gleich viele Durchgänge mit identischen Cue-Informationen in die Auswertung ein-

fließen⁵. In den ersten 20 Runden setzten die Versuchspersonen in der statistischen Umwelt T2P1 häufiger zweimal hintereinander auf dieselbe Alternative: Während das in T1P1 durchschnittlich 5.8 mal (SD= 1.7) der Fall war, ist dieser Wert in T2P1 mit 7.3 (SD= 3.0) eindeutig höher ($t = -2.4$; $p < .05$).

Das Entscheidungsverhalten ist vor allem auch im Hinblick auf die Häufigkeit interessant, mit der im Experiment die Alternative X (linke Seite auf dem Bildschirm) und die Alternative Y (rechte Seite) gewählt wurden. Ohne weitere Anhaltspunkte fingen die Versuchspersonen gewohnheitsmässig eher links an zu testen. Da nach einem ersten positiven Test in vielen Fällen bereits entschieden wurde, führt dies in der Summe dazu, dass bei unverbundenen Einzelereignissen T1P1 die Alternative X ($M=7.6$, $SD= 1.8$) signifikant häufiger gewählt als die Alternative Y ($M=5.0$, $SD=2.0$; $t = 4.0$; $p < .001$). In T2P1 hingegen ist dieser positionsbedingte Asymmetrieeffekt nicht zu beobachten: Die Alternative X ($M=6.3$, $SD=2.3$) wird gleich häufig gewählt wie die Alternative Y ($M=6.8$, $SD=3.1$; $t = -0.5$; $p = .61$). Offenbar reagieren die Versuchspersonen in der Umwelt der verbundenen Ereignisse stärker auf die jeweils vorausgehenden Ereignisse und nicht so sehr auf die Position der Alternative auf dem Bildschirm.

Diskussion des Vergleichs (between subjects)

Um die Wahrnehmung der Unterschiede in den statistischen Umwelten und das daran angepasste Verhalten (was sich in der Verhaltensmessung nicht von der Wahrnehmung trennen lässt) zu untersuchen, haben wir das Verhalten von zwei Versuchspersonengruppen in den zwei unterschiedlichen statistischen Umwelten miteinander verglichen.

Die Versuchspersonengruppen unterscheiden sich in der prognostizierten Richtung bei der Häufigkeit durchgeführter Tests, indem sie die logisch verbundenen Ereignisse zu nutzen wissen und vor allem positive Informationen aus der Vergangenheit berücksichtigen. Einige der Versuchspersonen benutzen diese, um in grosser Zahl ohne Durchführung neuer Tests zu entscheiden. Das führt zu einer erhöhten Anwendung der NNRStR, aber nicht zwingend zu einer erhöhten Summe von ORStR und NNRStR. Das bedeutet, dass die Reduktion der Informationssuche vor allem durch Entscheidungen ohne Informationkauf zustande kommt und weniger durch eine Verschiebung der MRStR in Richtung ORStR. Gleichzeitig dokumentieren die Versuchspersonen durch ihr Entscheidungsverhalten in T2P1, dass sie auf die vorauslaufenden Ereignisse in der Wahlwiederholung eingehen und auch nicht der spontanen Asymmetrie in der Bevorzugung der links stehenden Alternative aus T1P1 erliegen. Diese Asymmetrie der Alternativenwahl schien zuerst ein ungewollter Nebeneffekt des Versuchsaufbaus, wird hier nun aber zu einem Merkmal der adaptiven Informationssuche, indem sich diese Asymmetrie bei den verbundenen Ereignissen plötzlich aufhebt. Aus diesem Grund wird die Alternativenwahl in den weiteren Auswertungen als zusätzliche Fragestellung untersucht.

Insgesamt fallen die Effekte zwar eher geringer aus als bei einer perfekten strategischen Ausnutzung der statistischen Information zu erwarten wäre, aber dennoch zeigt sich in gewissen Mittelwertsunterschieden ein adaptives Verhalten an die beiden unterschiedlichen statistischen Umwelten.

⁵ Drei Personen haben jedoch weniger als 20 Runden absolviert: in T1P1 hat eine Person nur 17 Runden absolviert, in T2P1 haben zwei Personen 16 Runden und eine Person 17 Runden absolviert.

Detaillierte Resultate der Adaptation an den Wechsel der statistischen Umwelt (within subjects)

Für die Auswertung der Adaptation beim Wechsel der Umwelt wurden bei den inferenzstatistischen Auswertungen t-Tests für abhängige Stichproben mit zweiseitiger Hypothesenprüfung durchgeführt ($\alpha = .05$).

a) Resultate T1P1 - T2P2

In diesem Auswerteschritt wird die erste Personengruppe in ihrem Wechsel von einer Umwelt von unverbundenen statischen Ereignissen T1 zu einer Umwelt von verbundenen, dynamischen Ereignissen T2 untersucht. Mit der Annahme von eindeutig adaptivem Verhalten an die statistische Umwelt sollten sich die bisher gefundenen Effekte der unabhängigen Stichproben replizieren lassen.

Stoppverhalten

Anzahl durchgeführter Tests

In T1P1 wurden im Mittel 2.8 Cues (SD= 1.2) pro Runde gesucht, in T2P2 nur 2.4 Cues (SD= 1.2). Dieser Unterschied erweist sich nicht nur als signifikant ($t= 4.1$; $p < .001$), sondern er repliziert auch die gefundene Tendenz aus der Auswertung der unabhängigen Stichproben. Die Anzahl der gesuchten Cues über alle Runden gesehen unterscheidet sich nicht mehr signifikant zwischen den beiden statistischen Umwelten (T1P1 $M = 69.2$ Cues, $SD = 21.6$ Cues; T2P2 $M = 69.8$ Cues, $SD = 32.0$ Cues; $t= - 0.2$; $p = .83$).

Stoppregel

Wie bereits bei der Auswertung der unabhängigen Stichproben berichtet, stoppten die Versuchspersonen sowohl in T1P1 als auch in T2P1 etwas häufiger gemäss einer ORStR als gemäss einer MRStR (dieser Unterschied war jedoch für keine der beiden Umwelten signifikant). In T2P2 bleibt dieser Trend erhalten: Es wurde häufiger anhand einer Information entschieden ($M= 14.5$, $SD= 11.5$) als anhand mehreren Informationen ($M=12.7$, $SD= 10.8$), jedoch ist auch hier der Unterschied nicht signifikant ($t= 0.5$; $p=.64$).

Ebenfalls replizieren lässt sich der Befund, dass es im Vergleich zur ersten Bedingung T1P1 in der zweiten Bedingung T2P2 signifikant mehr Fälle gibt ($t= - 2.5$; $p < .05$), in denen überhaupt keine Informationssuche stattfindet (T1P1: $M=0.33$, $SD=0.6$ und T2P2: $M= 4.0$, $SD= 7.9$). Die Summe der NNRStR und der ORStR zeigt ein adaptives Verhalten der Versuchspersonen beim Wechsel gemäss Hypothese: In der Umwelt der verbundenen Ereignisse T2P2 wird signifikant häufiger mit keiner oder einer Information entschieden ($M= 18.5$, $SD= 13.9$) als in der Umwelt der Einzelereignisse T1P1 ($M=14.1$, $SD=12.6$; $t=-4.8$, $p<.001$).

Um zu sehen, wie die Personen individuell auf den Wechsel der Umwelt reagierten und ob sie ihre Stoppregel bei einer neuen Umwelt änderten, sind die intraindividuellen Daten genauer auszuwerten. Dazu lässt sich jede Versuchsperson einem dominanten „Stopping-Typ“ zuordnen, indem man die Häufigkeit von NNRStR, ORStR und MRStR über alle Runden bei dieser Person auszählt und die Person gemäss ihrer Majoritätsregel (mehr als die Hälfte aller Stoppregeln) klassifiziert (siehe die fett markierten Fälle in Tabelle 36).

Tabelle 36: Anzahl Stoppverhalten, welche bei den 30 Versuchspersonen in den beiden Bedingungen gemäss der No New Reason Stopping Rule (NNRStR), One Reason Stopping Rule (ORStR) und More Reason Stopping Rule (MRStR) klassifiziert werden. Die Majoritätsregel ist **fett markiert**.

T1P1			T2P2		
NNRStR	ORStR	MRStR	NNRStR	ORStR	MRStR
2	45	0	24	26	1
0	29	0	22	13	0
0	28	0	2	29	0
0	28	0	22	7	0
1	28	0	0	29	1
2	27	0	1	29	0
0	26	3	0	31	1
0	26	2	0	34	2
0	26	2	5	28	2
0	24	3	0	28	3
0	21	7	24	4	7
0	18	9	0	28	3
0	15	12	11	18	7
0	15	11	6	18	8
0	7	18	0	23	11
0	12	17	0	19	14
0	9	20	0	12	21
0	1	21	0	12	19
0	4	23	0	11	21
0	7	15	2	9	11
1	0	21	0	9	21
1	5	18	0	8	17
0	1	23	0	6	28
1	7	16	0	2	25
0	1	21	0	1	19
0	0	24	0	1	28
0	3	24	0	0	30
0	1	26	0	0	29
1	0	21	0	0	29
1	0	16	1	0	22

In der Bedingung T1P1 wendeten 14 Personen in mehr als der Hälfte der Fälle eine ORStR an und 16 Personen eine MRStR. In T2P2 wählten 12 Personen in mehr als der Hälfte der Fälle eine ORStR, 13 Personen stoppten gemäss MRStR und zwei Personen benutzten verschiedene Stoppregeln (Mix-Strategie). Drei Personen können in der Umwelt der verbundenen Ereignisse einer Strategie zugeteilt werden, anhand derer sie häufig ohne aktuelle Informationen auskamen (NNRStR).

77% der Versuchspersonen (23 von 30) verwendeten in beiden statistischen Umwelten dieselbe dominante Stoppstrategie (10 ORStR, 13 MRStR). Anhand von individuellen Differenzbeträgen zwischen der Anzahl ORStR und MRStR (ORStR minus MRStR) kann die Eindeutigkeit einer Strategiepräferenz berechnet werden. Ein niedriger Differenzbetrag (z.B. 3) bedeutet, dass die Versuchsperson keine eindeutige Präferenz für eine der beiden Strategien hat, ein hoher Wert (z.B. 45) deutet auf eine eindeutige Strategiepräferenz hin. Über alle Runden unterscheiden sich diese Differenzbeträge zwischen den beiden Bedingungen nicht ($t=0.3$; $p=.79$): In T1P1 beträgt die Differenz im Mittel 19.1 (SD= 9.2) und in T2P2 18.6 (SD= 9.7).

Obwohl sich die Wahl einer dominanten Stopppregel somit intraindividuell als recht stabil erweist und daher auch den Begriff des „Stopping-Typs“ gerechtfertigt ist, zeigen sich doch einzelne adaptive Anpassungen an die neue statistische Umwelt. Von den 14 One-Reason Decision Maker passten sieben Personen ihre Informationssuche an die Umwelt der verbundenen Ereignisse T2P2 an, indem sie häufig (in fünf oder mehr Fällen) eine NNRStR verwendeten. Die More-Reason Decision Maker entschieden sich in der neuen Umwelt T2P2 zwar kaum je ohne aktuelle Information, aber dennoch vertrauten sechs von sechzehn (inklusive zwei Strategiewechsler) auf vergangene Informationen, indem sie sich häufiger als in der vorherigen Umwelt (in mindestens zusätzlichen 5 Fällen) auch nur mit einer einzigen aktuellen Information entschieden.

Entscheidungsverhalten

Die Versuchspersonen setzten in den ersten 20 Runden in der statistischen Umwelt der verbundenen Ereignisse T2P2 häufiger zweimal hintereinander auf dieselbe Alternative: Während das in T1P1 durchschnittlich 5.8 mal ($SD = 1.7$) der Fall war, ist dieser Wert in T2P2 mit 7.4 ($SD = 2.8$) signifikant höher ($t = -3.0$; $p < .01$) und repliziert somit den Befund aus der Auswertung der unabhängigen Stichproben.

In der Auswertung der unabhängigen Stichproben hatten wir einen erheblichen Asymmetrie-Effekt zwischen Alternative X und Y in der Bedingung T1P1 gefunden. Dieser liess sich aus dem Versuchsaufbau erklären und zeigte sich nicht in T2P1. Dieser Effekt der aufgehobenen Asymmetrie verschwindet nun aber in T2P2 nicht, sondern die Asymmetrie verstärkt sich sogar noch: Die Alternative X wird in T2P2 wieder häufiger gewählt ($M = 7.4$, $SD = 1.7$) als die Alternative Y ($M = 4.2$, $SD = 2.5$; $t = 5.3$; $p < .001$). Bei dem Wechsel von T1P1 zu T2P2 zeigt sich das adaptive Verhalten, welches bei der unabhängigen Stichprobe in T2P1 beobachtet wurde, nicht mehr.

Diskussion des Vergleichs T1P1 - T2P2

In dieser Auswertung wurde untersucht, ob bei einem Wechsel von einer Umwelt mit unverbundenen Einzelereignissen in eine Umwelt mit statistisch verbundenen Ereignissen dieselben adaptiven Verhaltensunterschiede auftreten wie bei Personen, die jeweils nur eine dieser Umwelten bearbeiten. Hinsichtlich der Cuesuche können wir diese Frage mit einem klaren „Ja“ beantworten: Die Personen reduzieren nach dem Wechsel den Umfang ihrer Informationssuche in exakt demselben Ausmass an die Umwelt der verbundenen Ereignisse wie bei der Auswertung der unabhängigen Stichproben gefunden wurde. Dass diese Reduktion nur bei den Cuekäufen pro Runde, nicht aber bei den Cuekäufen über alle Runden zu finden ist, liegt an der grossen Differenz der Anzahl gespielter Runden.

Die Zuteilung der Personen zu einem dominanten „Stopping-Typ“ scheint sich in der Mehrzahl der Fälle (nämlich 77%) durch die im Experiment vorgenommene Änderung der Umwelt nicht beeinflussen zu lassen. Der Differenzbetrag beschreibt die Eindeutigkeit einer dominanten Stopppstrategie und unterstützt die bisherigen Befunde zu den relativ stabilen Anwendungen von Stopppregeln (Bröder, 2000a, 2000b; Zurbriggen, Christen, Hausmann, & Läge, 2008; Zurbriggen & Läge, 2007). Eine Betrachtung der intraindividuellen Daten zeigt aber, dass doch knapp die Hälfte der Personen (13 der 30) in der Umwelt der verbundenen Ereignisse

gehäuft ohne neue Information entscheiden und häufiger auf eine aktuelle Information statt auf mehrere Informationen vertrauen und sich somit in der Tendenz adaptiv verhalten.

In Bezug auf die Wahlwiederholung unterscheiden sich die Befunde der beiden T2-Bedingungen nicht, was darauf hindeutet, dass die Position der statistischen Umwelt auf diese Variable keinen Einfluss hat, dafür aber die statistische Umwelt. Die Personen passen ihr Verhalten in Bezug auf diese Variable an die statistische Umwelt an und verlassen sich in ihrer Wahl auf die vergangene Informationssuche. Sowohl in der statistischen Umwelt T1P1 wie auch in T2P2 ziehen die Versuchspersonen die Alternative X der Alternative Y vor. Es ist zu vermuten, dass in T2P2 die Aufhebung dieser Asymmetrie, welche in T2P1 beobachtet wurde, nicht mehr vorkommt, weil die Versuchspersonen in der Wahl der Alternative nicht adaptiv auf den Wechsel der statistischen Umwelt reagieren. Die in der ersten Umwelt gelernte Präferenz der erstplatzierten Alternative scheint in der Umwelt der verbundenen Ereignisse nicht mehr aufzulösen zu sein.

b) Resultate T2P1- T1P2

Im nächsten Auswerteschritt wird nun die zweite Personengruppe in ihrem Wechsel von einer Umwelt zur anderen betrachtet, und zwar in diesem Fall von T2P1 zu T1P2. Auch hier sollten sich – volle Adaptivität des Verhaltens annehmend – die in der Auswertung der unabhängigen Stichprobe gezeigten Effekte wiederholen.

Stoppverhalten

Anzahl durchgeführter Tests

In der Umwelt der verbundenen Ereignisse T2P1 und der Umwelt der unverbundenen Einzelereignisse T1P2 wurden im Mittel fast gleich viele Cues pro Runde gesucht (T2P1 $M = 2.3$, $SD = 1.4$, T1P2 $M = 2.2$, $SD = 1.2$, $t = -1.1$; $p = .29$). Auch unterscheidet sich die Anzahl der durchgeführten Tests über alle Runden nicht zwischen den beiden statistischen Umwelten (T2P1 $M = 57.2$, $SD = 23.8$; T1P2 $M = 63.2$, $SD = 27.9$; $t = 1.7$; $p = .09$).

Stoppregel

Wir untersuchen als nächstes wiederum die verschiedenen Stoppregeln: Einerseits die No New Reason Stopping Rule (NNRStR), die One Reason Stopping Rule (ORStR) und die More Reason Stopping Rule (MRStR). Eine optimale Adaptivität an die statistische Umwelt würde bei dieser Stichprobe bedeuten, dass sie sich in der Umwelt T1P2 nicht mehr ohne eine Information entscheiden sollten und häufiger anhand mehrerer Informationen (statt einer einzigen Information). Beim Wechsel von einer Umwelt mit verbundenen Ereignissen zu einer Umwelt von Einzelereignissen sollte daher die Summe von NNRStR und ORStR abnehmen. Genau das Gegenteil ist aber der Fall: in T2P1 wird im Mittel 16.7mal ($SD = 10.1$) anhand keiner oder einer Information entschieden, in T1P2 aber 20.0mal ($SD = 10.8$). Der Unterschied ist denn auch signifikant ($t = 2.189$, $p > .05$).

Es zeigt sich, dass in der statistisch unverbundenen Umwelt T1P2 die Personen nur etwas weniger auf die Informationssuche verzichteten (T2P1 $M = 3.3$, $SD = 5.7$; T1P2 $M = 1.8$, $SD = 4.2$; $t = 0.3$; $p = .10$) und häufiger eine ORStR anwendeten anstatt einer MRStR. Über alle Runden hinweg gesehen, entschieden sich in beiden Bedingungen etwas mehr Personen an-

hand einer Information (T2P1 M= 13.4, SD= 7.9; T1P2 M=18.2, SD= 10.9) als anhand mehrerer (T2P1 M=10.5, SD= 6.8; T1P2 M= 9.7, SD= 10.0). In T2P1 ist dieser Unterschied nicht signifikant ($t=1.22$; $p=.23$) und repliziert somit alle bisherigen Befunde. In der zweiten Bedingung T1P2 sieht das Bild nun aber etwas anders aus, denn der Unterschied wird signifikant ($t= 2.4$; $p< .05$). Das heisst, dass auf Mittelwertsebene zu häufig nach dem Wechsel gemäss einer ORStR und wenig gemäss einer MRStR gestoppt wird. Um zu sehen, wie die einzelnen Personen auf den Umweltwechsel mit ihrer Stoppregel reagierten, lohnt sich auch hier wieder ein Blick auf die individuellen Daten in Tabelle 37.

Tabelle 37: Anzahl Fälle, welche bei den 30 Versuchspersonen in den beiden Bedingungen gemäss der No New Reason Stopping Rule (NNRStR), One Reason Stopping Rule (ORStR) und More Reason Stopping Rule (MRStR) klassifiziert werden. Die Majoritätsregel ist **fett markiert**.

T2P1			T1P2		
NNRStR	ORStR	MRStR	NNRStR	ORStR	MRStR
18	10	0	18	5	0
21	9	1	0	30	0
13	30	3	0	30	1
0	27	1	0	31	0
0	26	0	0	25	0
5	23	4	2	32	0
0	20	10	0	28	3
0	20	10	0	11	22
0	18	9	1	19	9
0	17	10	0	31	0
3	17	6	6	27	0
5	17	2	1	10	19
10	17	1	14	13	0
8	16	6	2	27	0
0	15	12	0	31	0
1	1	23	0	1	23
0	2	21	0	12	19
0	17	19	0	16	20
2	9	18	0	11	24
0	0	17	0	0	27
0	7	17	0	30	1
0	10	16	0	9	14
0	13	16	1	19	11
0	1	15	1	0	25
1	0	15	1	0	19
0	12	14	0	23	11
0	13	14	0	19	11
0	13	14	0	31	0
2	14	14	0	12	24
9	9	6	7	14	8

Teilt man die Versuchspersonen gemäss der Majoritätsregel einem „Stopping-Typ“ zu, findet man in der ersten Bedingung T2P1 zwei Personen, die meist ohne neue Information entschieden (NNRStR), 13 Personen, die häufig die Informationssuche gemäss ORStR stoppten und 13 Personen, die überzufällig viel MRStRs anwendeten. Zwei Personen verwendeten keine der drei Strategien überzufällig häufig. In der Umwelt der unverbundenen Ereignissen, also der zweiten Bedingung T1P2, stoppten 16 Personen überzufällig häufig gemäss einer ORStR, 11 Personen stoppten gemäss MRStR, zwei Personen entschieden sich überzufällig häufig ohne Information und eine Person verwendete alle Stoppregeln.

Die Subgruppen der „Stopping-Typen“ in der Umwelt der verbundenen Ereignisse T2P1 entsprechen den Gruppengrößen, welche in derselben Umwelt T2P2 gefunden wurden. In der Umwelt der unverbundenen Einzelereignisse T1P2 ist dies jedoch nicht der Fall. Zeigten sich in T1P1 ungefähr zwei gleich grosse Subgruppen (14 ORStR-Typen und 16 MRStR-Typen), findet man nur in derselben Umwelt T1P2 eindeutig mehr Personen die klar ein *One Reason Decision Making* betreiben (16 ORStR-Typen und 11 MRStR-Typen).

Wie stabil die Strategien bei diesem Umweltwechsel sind, zeigt sich auch hier: Bei insgesamt 63% der Versuchspersonen ist die Klassifikation in beiden Umwelten dieselbe (10 ORStR, 8 MRStR, 1NNRStR). Fünf Personen passten ihre Stoppregel an die neue Umwelt an, indem sie von einer dominanten MRStR zu einer dominanten ORStR wechselten, zwei Personen verhielten sich genau konträr und wechselten von einer dominanten ORStR zu einer eindeutigen MRStR. Eine Person adaptierte ihr Verhalten, indem sie von häufigen Raten zu vermehrten ORStRs wechselte und eine Person tat genau das Umgekehrte. Zudem wechselte eine Person von keiner spezifischen Zuteilung zur dominanten Verwendung von MRStR, und eine Person bevorzugte keine der beiden Strategien in beiden Umwelten.

Berechnet man den Differenzbetrag zwischen den beiden verwendeten Stoppregeln pro Person (ORStR minus MRStR), findet man in T2P1 einen niedrigeren Wert ($M = 11.0$, $SD = 1.4$) als in T1P2 ($M = 18.8$, $SD = 10.1$; $t = 4.1$, $p < .001$). Die Personen scheinen zuerst, wenn mit der Umwelt der verbundenen Ereignisse konfrontiert, beide Stoppstategien anzuwenden, dann aber bei der Umwelt mit unverbundenen Ereignissen sich auf eine Strategie festzulegen, und zwar, wie oben gezeigt wurde, entscheiden sie sich häufiger anhand einer einzigen Information.

Gemäss unserer Hypothese müssten die Versuchspersonen in T1P2 seltener (d.h. in weniger als fünf Fällen) im Bezug zu zuvor ohne aktuelle Information die Informationssuche stoppen und dafür eine kaufen. Es passten jedoch nur drei Personen in diesem Sinne ihre Informationssuche an die neue statistische Umwelt an. Zudem müssten die Personen, welche sich in T2P1 auf eine Information verliessen, in T1P2 häufiger mehrer Informationen anschauen. Es verändern aber nur sechs Personen ihre Präferenz von einer Information zu mehreren Informationen. Diejenigen Personen, die in der Umwelt die Struktur der Abhängigkeit effizient auszuschöpfen wussten, scheinen Gefahr zu laufen, bei einem Wechsel bei dieser Strategie haften zu bleiben.

Entscheidungsverhalten

In T1P2 wurde von 20 Runden durchschnittlich 6.5 mal ($SD = 3.6$) zweimal hintereinander auf dieselbe Alternative gesetzt, in T2P1 ist dieser Wert mit 7.3 ($SD = 3.0$) etwas höher, jedoch nicht signifikant ($t = -1.5$; $p = .15$). Die verminderte Wiederwahl in der Umwelt der Einzelereignisse lässt sich somit von der Tendenz her wieder finden, in T1P1 waren es 5.8 mal, aber das Verhalten in T1P2 ist nicht adaptiv genug, als dass es signifikant würde. Hingegen adaptieren die Versuchspersonen bei der Alternativenwahl in die gleiche Richtung wie die unabhängige Stichprobe: Die beobachtete Asymmetrie der Alternativenwahl, d.h. Präferenz der Alternative X ($M = 7.5$, $SD = 2.7$) gegenüber der Alternative Y ($M = 5.0$, $SD = 3.4$) stellt sich in der Umwelt der unverbundenen Ereignisse nach einem Wechsel von T2P1 zu T1P2 wieder ein ($t = 2.5$; $p < .05$).

Diskussion des Vergleichs T2P1 - T1P2

Es muss an dieser Stelle vermutet werden, dass diese Aufgabe des Wechselns von einer statistischen Umwelt mit verbundenen Ereignissen in eine Umwelt mit statistisch unverbundenen Einzelereignissen für die Versuchspersonen schwieriger zu lösen ist als die umgekehrte Variante. Die Schwierigkeit resultiert daraus, dass die Versuchspersonen zuerst Informationen über eine Zeitkonstanz bekommen, welche ihnen danach in der zweiten Bedingung vorenthalten wird. Dennoch wurde in dieser Auswertung ebenfalls untersucht, ob bei diesem Wechsel dieselben adaptiven Verhaltensunterschiede auftreten wie bei Personen, die jeweils nur eine dieser beiden Umwelten bearbeiten. Diese Frage können wir in dieser Auswertung nur hinsichtlich des Entscheidungsverhaltens bejahen und auch dies nur bedingt.

Die Cuekäufe werden nicht adäquat an die neue statistische Umwelt angepasst und auch die Stopppregel nicht. Im Gegenteil die Personen bevorzugen nach einem Wechsel in der Umwelt der unverbundenen Einzelereignisse T1P2 eindeutig häufiger eine ORStR anstatt einer MRStR. Die vermehrte Informationssuche, welche in T1P1 gefunden wurde, wird nach dem Wechsel nicht repliziert. Obwohl dieses Verhalten als eine Nicht-Anpassung an die statistische Umwelt der Einzelereignisse interpretiert werden kann, „optimieren“ die Versuchspersonen statistisch gesehen ihre Informationssuche bei dem Wechsel auf nahezu zwei Informationen pro Durchgang.

Die einzigen replizierten Befunde sind jene der Häufigkeit einer Wiederwahl in der unverbundenen Umwelt und jene zur Asymmetrie in der Alternativenwahl. Die Reduktion an Wahlwiederholung in der Umwelt der Einzelereignisse ist insofern sinnvoll, als dass in dieser Umwelt keine Zeitkonstanz vorhanden ist. Zudem legen die Versuchspersonen in der Umwelt der Einzelereignisse, nach einem Wechsel, wieder ein grösseres Gewicht auf die links positionierte Alternative und wählen seltener dieselbe Alternative zweimal hintereinander.

Im Grossen und Ganzen lassen sich die Befunde aus der Auswertung der unabhängigen Stichproben in dieser Auswertung nicht replizieren. Bei dem Wechsel von einer Umwelt zeitlich verbundener Ereignisse zu einer Umwelt mit unverbundenen Einzelereignissen verändert sich insbesondere das Stoppverhalten nicht in der angenommenen Weise, gewisse Aspekte des Entscheidungsverhaltens jedoch schon.

Diskussion der Adaptation an den Wechsel der statistischen Umwelt (within subjects)

In dieser Auswertung wurde der Frage nachgegangen, ob die in der ersten Auswertung gefundenen Unterschiede einfach Personengruppenunterschiede sind oder ob sich ein systematischer Unterschied in der Adaptivität an die zwei statistischen Umwelten finden lässt. Leider lassen sich die in der Auswertung der unabhängigen Stichproben gefundenen Unterschiede nur bedingt replizieren. Auffällig ist, dass sich eindeutig andere Verhaltensmuster zeigen, je nach dem mit welcher statistischen Umwelt das Experiment gestartet wird. Starten die Versuchspersonen mit den logisch unverbundenen Ereignissen, dann reduzieren sie beim Wechsel erwartungsgemäss die Informationssuche und verhalten sich somit adaptiv. Diese Adaption der Cuekaufhäufigkeit lässt sich beim Wechsel von einer verbundenen Umwelt zur unverbundenen nicht mehr finden. Weshalb, so stellt sich die Frage, werden in T1P2 so wenige Infor-

mationen gekauft oder müsste man eher fragen: Weshalb werden in T1P1 so viele Informationen gekauft? Welches Verhalten ist nun adaptiv? Das ACT-R Modell zu adaptivem Verhalten postuliert eine Explorationsphase (Taatgen, Lebiere, & Anderson, 2006) und könnte ein Ansatzpunkt zur Erklärung unserer experimentellen Befunde in T1P1 sein. Womöglich suchen die Personen in der ersten Bedingung so viele Informationen, weil sie die Umwelt zuerst explorieren möchten.

Ebenfalls unterschiedliche Verhaltensmuster zeigen sich bei der Anwendung der beiden Stoppregeln ORStR und MRStR. Der Differenzbetrag sagt etwas über die Eindeutigkeit einer dominanten Stoppstrategie aus: Je nach dem mit welcher statistischen Umwelt begonnen wird, ist diese Eindeutigkeit unterschiedlich. Wenn die Versuchspersonen das Experiment mit der Umwelt der verbundenen Ereignisse beginnen, dann haben sie Mühe, eine dominante Stoppstrategie zu entwickeln. Beginnen die Versuchspersonen das Experiment mit der Umwelt der logisch unverbundenen Einzelereignisse, bevorzugen sie eine klare Stoppstrategie. Betrachtet man die Daten auf Personenebene, so zeigt sich, dass 42 der 60 Personen in beiden statistischen Umwelten mehrheitlich dieselbe Stoppstrategie wählen, d.h. nicht sonderlich stark auf einen Wechsel der Umwelt reagieren. Adaptiv verhalten sich einige One-Reason Decision Maker, indem sie in der Umwelt der verbundenen Ereignisse auch mal ohne eine aktuelle Information entscheiden und einige More-Reason Decision Maker, indem sie sich auch ab und zu mit einer einzigen Information zufriedengeben. Diese Anpassung lässt sich beim Wechsel von einer verbundenen Umwelt zu einer unverbundenen Umwelt nicht im gleichen Ausmass finden. Da sollte nach dem Wechsel signifikant weniger NNRStR und mehr MRStR stattfinden, stattdessen ist - abgesehen von zwei Strategiewechslern - das Gegenteil zu beobachten: Die Versuchspersonen legen sich nach einem Wechsel noch mehr auf ein *One Reason Decision Making* fest.

Ein eindeutig adaptives Verhalten lässt sich einzig bei den Wahlwiederholungen beobachten und bei beiden Wechseln auch replizieren. Im Unterschied zur Umwelt von unverbundenen Ereignissen zeigen die Versuchspersonen in der Umwelt der verbundenen Ereignisse häufiger Wahlwiederholung als in der Umwelt der unverbundenen Einzelereignisse, und dies unabhängig von der Position der statistischen Umwelt im Experiment. Dieses Resultat kann in die bisherigen Befunde zu Routineverhalten im zeitlichen Kontext (siehe z.B. Betsch, Haberstroh, Glöckner, Haar, & Fiedler, 2001 oder Bröder, 2006) eingereiht werden. Die ansonsten durchgehende Präferenz der links positionierten Alternative X wird aufgehoben, wenn die Ereignisse miteinander verbunden (T2) sind und T2 an erster Stelle im Experiment steht. Dieses adaptive Verhalten lässt sich aber beim Wechsel zu einer Umwelt von verbundenen Ereignissen (T1P1 - T2P2) nicht replizieren.

Diese Mehrzahl von Befunden zur Nicht-Adaptivität im Bezug auf die Informationssuche und das Stoppverhalten stützen die Befunde von Rieskamp, wo die Versuchspersonen ebenfalls eine langsame Adaption an eine neue Situation zeigten. Auch Rakow et al. (2005) fanden in ihren Befunden ein träges adaptives Verhalten beim Umweltwechsel und erklären es so, dass die Personen anfangs vermutlich ein Aspirationslevel setzen und dann nach einem Umweltwechsel dabei bleiben.

Schlussbemerkung

Bei der Auswertung betreffend der Adaptivität an eine Umwelt zeigen sich einige Verhaltensanpassungen in die von uns erwartete Richtung: Es werden in der Umwelt der verbundenen Ereignisse weniger Tests durchgeführt, die Personen entscheiden sich ab und zu ohne neue Informationen und sie orientieren sich in ihren Entscheidungen stärker an vorausgehenden Entscheidungen als in der Umwelt der Einzelereignisse. Das Entscheiden ohne aktuelle Informationen und die verstärkte Wahlwiederholung bedeutet, dass die Versuchspersonen sich an den vorauslaufenden Informationen orientieren und sie grundsätzlich den Unterschied der beiden statistischen Bedingungen verstanden haben. In Bezug auf die Stopppregel verhalten sich die Personen nicht so wie wir dies in den Hypothesen formuliert haben, z.B. passen sie die zwei Stopppregeln ORStR und MRStR nicht in der erwarteten Weise an die statistische Umwelt an.

Wenn nun die Versuchspersonen von einer statistischen Umwelt in die andere wechseln, dann zeigt sich eine eher geringe Verhaltensanpassung an die neue statistische Umwelt, da die Auswirkungen der zuvor bearbeiteten statistischen Umwelt gross sind. Einzig bei den Wahlwiederholungen und der Wahl, ohne aktuelle Information eine Entscheidung zu fällen, ist eine Verhaltensanpassung an die statistische Umwelt, fast gänzlich unabhängig von der vorherigen Bedingung, zu beobachten. In allen anderen Fällen, in denen sich die Versuchspersonen nicht unmittelbar auf die neue Umwelt einstellen, eröffnet sich die Frage nach dem Einfluss der Position der Bedingung und der Interaktion der beiden Faktoren „statistische Umwelt“ und „Position“. Der Einfluss der Position zeigt sich statistisch bei der Summe von NNRStR und ORStR. Einen Interaktionseffekt finden wir bei der asymmetrischen Wahl der Alternative. Es zeigt sich somit bei der gesamten Auswertung, dass es schwierig auszumachen ist, was ein Effekt der statistischen Umwelt und was ein Effekt der Position dieser statistischen Umwelt im Experiment ist. Teilweise entsprechen die Resultate zwar den Annahmen, die wir betreffend dem adaptiven Verhalten an eine statistische Umwelt gemacht haben, aber dann wiederum lassen sich diese Annahmen bei einem Wechsel der Umwelt bestätigen, bei dem anderen Wechsel wiederum nicht.

Die Befunde fügen ein paar neue Fragezeichen ans Konzept der Adaptiven Toolbox an, gemäss derer Menschen verschiedene kognitive Werkzeuge je nach Umwelt adaptiv auswählen und anwenden. Bröder und Newell kritisieren an der adaptiven Toolbox-Theorie u.a., dass der Selektionsprozess dieser kognitiven Werkzeuge ressourcenaufwändig zu sein scheint und dieser Aufwand aber in der Theorie keinen Platz fände (Bröder & Newell 2008). Unsere Befunde ergänzen diese Kritik und lassen zudem vermuten, dass dieser Selektionsprozess einerseits durch einen Explorationseffekt in einer ersten statistischen Umwelt, andererseits durch einen Trägheitseffekt in der nächsten statistischen Umwelt beeinflusst wird.

Zusammenfassend lässt sich sagen, dass Menschen zwischen den beiden Situationsklassen von verbundenen und unverbundenen Ereignissen allerdings sehr wohl unterscheiden können. Vor allem die Berücksichtigung vergangener Informationen und die Wiederwahl der Alternativen belegt dies. Versetzt man die Versuchspersonen jedoch anfangs in eine gewisse statistische Situationsklasse, haben sie Mühe, diese zu wechseln und bleiben in ihrem Verhalten verhaftet. Will man das Konzept der Adaptiven Toolbox möglichst uneingeschränkt aufrecht

erhalten, so tut man sicherlich nicht schelcht daran, Adaptation und Routine als antagonistische Verhaltenstendenzen herauszuarbeiten und deren Wirkweise im Prozess der Entscheidungs-genese umweltabhängig zu studieren.

Literatur

- Betsch, T., Haberstroh, S., Glöckner, A., Haar, T., & Fiedler, K. (2001). The effects of routine strength on adaptation an information search in recurrent decision making. *Organizational Behavior and Human Decision Processes*, 84(1), 23-53.
- Bröder, A. (2000a). Assessing the empirical validity of the "Take The Best"-heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1332-1346.
- Bröder, A. (2000b). "Take the best - Ignore the rest" Wann entscheiden Menschen begrenzt rational? Lengerich: Pabst.
- Bröder, A. (2005). Entscheiden mit der "adaptiven Werkzeugkiste". Ein empirisches Forschungsprogramm. Lengerich: Pabst Science Publishers.
- Bröder, A., & Schiffer, S. (2003a). Bayesian strategy assessment in multi-attribute decision making. *Journal of Behavioral Decision Making* 16, 193-213.
- Bröder, A., & Schiffer, S. (2003b). Take The Best versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General*, 132, 277-293.
- Bröder, A., & Schiffer, S. (2006). Adaptive flexibility and maladaptive routines in selecting fast and frugal decision strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(4), 904-918.
- Bröder, A., & Newell, B. (2008). Challenging some common beliefs: Empirical work within the adaptive toolbox metaphor. *Judgement and Decision Making*, 3(3), 205-214.
- Christen, S., Zurbriggen, S., Hausmann, D., & Läge, D. (2007). Trinkwasser- und Börsenszenario: Aufbau, Ablauf und technische Spezifikation einer Experimentalumgebung zur entscheidungsvorbereitenden Informationssuche in zeitlichen Verläufen. AKZ-Forschungsbericht 43, Zürich: Angewandte Kognitionspsychologie.
- Cosmides, L., & Tooby, J. (1992). The adapted mind: Evolutionary psychology and the generation of culture. New York: Oxford University Press.
- Czerlinski, J., Gigerenzer, G., & Goldstein, D. (1999). How good are simple heuristics? In G. Gigerenzer, P.M. Todd & The ABC Research Group (Eds.), *Simple Heuristics That Make Us Smart* (pp. 97-118). New York: Oxford University Press.
- Fasolo, B., Misuraca, R., & McClelland, G. H. (2003). Individual differences in adaptive choice strategies. *Research in Economics*, 57(3), 219-233.
- Fu, W.-T., & Gray, W. D. (2006). Suboptimal Tradeoffs in information seeking. *Cognitive Psychology*, 52(2), 195-242.
- Gigerenzer, G., & Goldstein, D. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650-669.
- Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The Take The Best heuristic. In G. Gigerenzer, P.M. Todd & The ABC Research Group (Eds.), *Simple Heuristics That Make Us Smart* (pp. 75-95). New York: Oxford University Press.
- Gigerenzer, G., & Todd, P. (1999). Fast and Frugal Heuristics - The Adaptive Toolbox. In G. Gigerenzer, P.M. Todd & The ABC Research Group (Eds.), *Simple Heuristics That Make Us Smart* (pp. 3-34). New York: Oxford University Press.
- Goldstein, D., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109, 75-90.
- Hogarth, R. M., & Karelaia, N. (2006). "Take-the-Best" and other simple strategies: Why and when they work "well" with binary cues. *Theory and Decision*, 61, 205-249.
- Läge, D., & Hausmann, D. (2007). Der Einfluss der Informationssuche auf die Entscheidung. AKZ-Forschungsbericht 4, Zürich: Angewandte Kognitionspsychologie.
- Newell, B., Rakow, T., Weston, N., & Shanks, D. (2004). Search strategies in decision making: The success of 'success'. *Journal of Behavioral Decision Making*, 17, 117-137.
- Newell, B., & Shanks, D. (2003). Take the best or look at the rest? Factors influencing one-reason decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 53-65.
- Newell, B., Weston, N., & Shanks, D. (2003). Empirical tests of a fast and frugal heuristic : not everyone "takes-the-best". *Organizational Behavior and Human Decision Processes*, 91, 82-96.
- Pfeifer, R., & Scheier, C. (1999). *Understanding Intelligence*. Cambridge: MIT Press.
- Pinker, S. (1994/1996). *The Language Instinct*. New York: William Morrow and Company. Deutsch: *Der Sprachinstinkt*. München: Kindler.

- Rakow, T., Newell, B., Fayers, K., & Hersby, M. (2005). Evaluating three criteria for establishing cue-search hierarchies in inferential judgment. *Journal of Experimental Psychology / Learning, Memory & Cognition*, 31(5), 1088-1104.
- Rieskamp, J. (2006). Perspectives of probabilistic inferences: reinforcement learning and an adaptive network compared. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(6), 1355-1370.
- Rieskamp, J. (2008). The importance of learning when making inferences. *Judgment and Decision Making*, 3, 261-277.
- Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics and how can we tell? In G. Gigerenzer, P.M. Todd & The ABC Research Group (Eds.), *Simple Heuristics That Make Us Smart* (pp. 141-168). New York: Oxford University Press.
- Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135, 207-236.
- Taatgen, N. A., Lebiere, C., & Anderson, J. R. (2006). Modeling paradigms in ACT-R. In R. Sun (Ed.), *Sognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation* (pp. 29-52): Cambridge University Press.
- Zurbriggen, S., Christen, S., Hausmann, D., & Läge, D. (2008). Entscheiden in zeitlich dynamischen Umwelten: Suchen und Stoppen mit unterschiedlich stabilen Cues. *AKZ-Forschungsbericht 44*, Zürich: Angewandte Kognitionspsychologie.
- Zurbriggen, S., Hausmann, D., Christen, S., & Läge, D. (2007). Die zeitliche Dynamik von Umwelten: Heuristisches Wahlverhalten am Beispiel eines Trinkwasser-Szenarios mit und ohne Zeitachse. *AKZ-Forschungsbericht 41*, Zürich: Angewandte Kognitionspsychologie.
- Zurbriggen, S., & Läge, D. (2007). Ressource Zeit statt Geld: Ein Vergleich zweier Szenarien zur Entscheidungsforschung. *AKZ-Forschungsbericht 45*, Zürich: Angewandte Kognitionspsychologie.

KRITISCHE BEMERKUNGEN

Kapitel 7: Die Präsentation von Cue-Gütemassen im Information Board Setting: Validität, Diskriminationsrate und Stabilitätsrate

Einleitung

Ursprünglich definierten Gigerenzer et al. (1991) in ihrem probabilistischen mentalen Model (PMM) die Cue-Validitäten im Bezug auf probabilistische Informationen im Gedächtnis: „A PMM connects the specific structure of the task with a probability structure of a corresponding natural environment (stored in long-term memory)” (1991, S.507). Diesbezüglich ist die Informationssuche in diesem Paradigma als ein Informationsabruf im Gedächtnis gemäss den gelernten Cue-Validitäten konzipiert. Eine solche Theorie lässt sich zwar relativ einfach anhand Simulationsstudien überprüfen, doch überprüfbare Verhaltensdaten zu generieren ist schwierig (siehe dazu die Ausführungen von Bröder zum Identifizierungsproblem, 2000b). Verschiedene methodische Möglichkeiten bieten sich an, dieses Problem zu lösen: z.B. kann anhand von Zeitmessungen der Gedächtnisabruf operationalisiert werden und so Schlussfolgerungen über die Informationssuche im Gedächtnis getroffen werden (z.B. Bergert & Nosofsky, 2007; Bröder & Gaissmaier, 2007). Eine andere Vorgehensweise beruht auf einer „Externalisierung“ der Informationssuche, indem die Informationssuche anhand eines Information Boards (basierend auf der Idee von Payne, 1976) überprüft wird. Dass der letztere Ansatz nicht dem PMM-Paradigma entspräche, kritisieren Gigerenzer und Todd (1999): “We refer to this type of task as inference from givens, as opposed to inference from memory or inference from the external environment, both of which require search. [...] Experiments in which search is obviated are unsuitable for testing models of ecological and bounded rationality that rely on limited information search as a central component” (1999, S. 23). Diese Kritik war gerechtfertigt, solange alle Informationen auf dem Information Board vorgegeben wurden und somit keine aktive und ressourcenaufwändige Informationssuche möglich war. Um den theoretischen Annahmen der Bounded Rationality nachzukommen, werden aber seit längerem die Kosten der Informationssuche im Gedächtnis durch monetäre Kosten (z.B. Bröder, 2000a; Bröder, 2003; Hausmann & Läge, 2005; Läge, Hausmann, Christen, & Daub, 2005; Newell & Shanks, 2003) oder durch zeitliche Kosten (z.B. anhand Zeitdruck Rieskamp & Hoffrage, 1999; oder anhand fiktiver Zeit Zurbriggen, Hausmann, Christen, & Läge, 2007) experimentell implementiert. Die Methodik des Information Boards bleibt damit eine *inference from givens*, aber eröffnet dennoch die Möglichkeit, die theoretischen Annahmen des PMM anhand von Verhaltensdaten zu überprüfen.

Nach einer Reihe von Lernexperimenten gehen nun verschiedene Autoren dazu über, unterschiedliche Cue-Gütemasse (wie Validität, Diskriminationsrate oder neu die Stabilitätsrate) auf dem Information Board numerisch anzugeben. Mit dieser Massnahme kann das Wissen über die Cue-Güte exakt vorgegeben werden, und somit können auch klare Schlussfolgerungen über die Anwendung der Gütemasse in Bezug auf die sequentielle Informationssuche getroffen werden. Ein Überblick soll zeigen, wie es zu dieser Praxis gekommen ist, um dann zu prüfen, ob dieses Vorgehen bei dem Gütemass der Cue-Stabilität überhaupt zulässig ist.

Die Validität von Cues

Um das PMM zu überprüfen und einer gewissen Cue-Reihenfolge in der Informationssuche folgen zu können, muss zuerst Wissen über die Güte dieser Cues vorhanden sein. Als erstes Cue-Gütemass wurde die Validität untersucht. Bröder formulierte das Cue-Learning Paradigma anfangs so, dass er in einer Lernphase die Cue-Ausprägungen lernen liess und dann in einer Inferenzphase die Personen anhand aller Informationen entscheiden liess (2000b). Der Autor argumentierte dabei gegen das Mitteilen von Cue-Validitäten: „Eine Information über die Validitäten setzt zunächst voraus, dass das formale Konzept der Validität den Vpn erklärt und von diesen auch verstanden wird und birgt demnach die Gefahr, dass dies nicht immer erreicht werden kann. Zudem würde diese explizite Aufklärung über statistische Konzepte den Vpn möglicherweise einen „statistischen“ Entscheidungsmodus nahelegen, also *Demand Characteristics* induzieren. Schwerwiegender ist der weitere Einwand, dass beide Methoden (sequenzielles Lernen vs. Mitteilen) möglicherweise zu unterschiedlichen internen Repräsentationen der Validität führen“ (2000b, S.91). Bröder fand in der Nachbefragung, in welcher er die Cue-Validitäten erläuterte und die Versuchspersonen die Validitäten der Cues numerisch einschätzen, dass die Versuchspersonen tendenziell hohe Validitäten unterschätzen und tiefe überschätzten. Numerisch waren die subjektiven und objektiven Validitäten somit nicht identisch, aber die Reihenfolge der subjektiven Validitäten entsprach der Reihenfolge der objektiven Validitäten (die Mittelwerte der Korrelationen lagen bei vier Experimenten jeweils über .96). Um die sequentiellen Suchregeln gemäss PMM zu überprüfen, folgte Bröder, dass einzig eine lineare Gleichheit der Validitäten nötig wäre: „Die Ergebnisse sprechen also insgesamt dafür, dass mit dem Cue-Learning eine für den Einsatz der untersuchten Entscheidungsregeln adäquate subjektive Repräsentation der Cue-Validitäten erreicht werden kann“ (2000b, S.99). In einem weiteren Experiment zu aktivem und passivem Lernen (Experiment 3, 2000b, S.136) zeigte der Autor, dass die gelernte subjektive Hierarchie der Cues durch passives Lernen fast genauso gut eingeschätzt werden konnte wie durch aktives Lernen (Mittelwerte von .95 bzw. .99). Dieser Befund führte dazu, dass Bröder in weiteren Experimenten die aktive Lernphase zuerst verkürzte (z.B. in Experiment 2, Bröder, 2000a) und sie dann später ganz aufhob. Im „Börsenspiel“ ging Bröder dann nämlich dazu über, den Personen das Konzept der Cue-Validitäten zu erklären und die Cues und ihre einzelnen Validitäten zu beschreiben (2000b, S.171) und die Personen diese anhand beliebig vieler Durchgänge erfahren zu lassen. Diese Lernmethode behielt der Autor bei den weiteren Experimenten (Bröder, 2005) zum Börsenspiel (Hierarchieangabe, welcher Cue der beste Prädiktor ist) und dem Immobilienfond (Angabe der Wichtigkeit der Cues) bei. Bröder kommentierte dies folgendermassen: „Im Unterschied zu den ersten beiden Experimenten wurden die Cue-Validitäten nicht mehr in einer aufwändigen Lernphase gelernt, sondern in den Instruktionen erläutert und mitgeteilt“ (S.64). In den rein gedächtnisbasierten Experimenten zum sogenannten „Kriminalspiel“ liess Bröder die Cue-Ausprägungen wieder ausführlicher lernen (Bröder, 2005; Bröder & Schiffer, 2003, 2006), wobei er den Versuchspersonen nach der Lernphase dennoch mitteilte, welche Indizien (Cues) am Tatort aufzufinden waren und mit welcher Wahrscheinlichkeit (=Validität) sie zum Täter passten (Bröder, 2005, S.199). Somit gab er die Hierarchie der Cue-Validitäten auch in diesen Experimenten bekannt: z.B. „It was made plausible by the cover story that the validity hierarchy of the evidence was blood type, cigarette brand, perfume, vehicle“ (Bröder & Schiffer, 2003, S.283). Die Vorgabe von den Rangreihen der Cue-

Validitäten war insofern sinnvoll, als dass Bröder hauptsächlich die Entscheidungsregel interessierte. Die Befunde sagen aber wenig darüber aus, wie der Autor selbst in einem aktuellen Artikel bemerkt (Bröder & Newell, 2008), anhand welcher Suchregeln Personen aktiv Informationen suchen.

Newell und Shanks (2003) übernahmen das Börsenszenario von Bröder, änderten aber die Lernphase der Cues. In 60 Trainingsdurchgängen wurden den Versuchspersonen jeweils das ganze Cue-Muster (d.h. Cues mit binären Ausprägungen für zwei Aktien) vorgegeben. Auf Basis dieser Informationen mussten die Versuchspersonen danach auf eine Aktie tippen, worauf sie ein Feedback erhielten. In der Testphase mussten die Personen dann aktiv Informationen kaufen, um sich für eine Aktie zu entscheiden. Die Personen lernten in dem ersten durchgeführten Experiment die Validitäten sehr schlecht, d.h. die Rangreihe der gekauften Cues entsprach zwar der objektiven Rangreihe im Trend (wobei der verwendete Jonckheere-Test nicht als sehr sensibler Test gilt), aber die Häufigkeiten der Cuekäufe in der aktiven Informationssuche unterschieden sich nur zwischen dem zweit- und drittvalidesten Cue. Dies führte dazu, dass die Autoren in dem zweiten und dritten Experiment sogenannte *hints* (Tipps) einschalteten. In der Mitte der Lernphase und direkt am Anfang der Testphase wurde den Versuchspersonen die korrekte Reihenfolge der Cue-Validitäten vorgegeben. Genau gleich verfahren Newell, Weston und Shanks (2003). Dieses Verfahren macht die Überprüfbarkeit des Lernens der Cue-Validität unmöglich, was insofern zu entschuldigen ist, als dass in letzterer Untersuchung vor allem die Stoppregel im Fokus stand.

Hausmann, Christen & Läge (2005) liessen ebenfalls Versuchspersonen Cue-Validitäten im „Millionärsquiz“ lernen. Es zeigte sich, dass die Cue-Validitäten mehr oder weniger ordinal richtig eingeschätzt wurden (das heisst 70% der Versuchspersonen erzielten eine Korrelation grösser als .80), gleichzeitig aber die absoluten Validitäten systematisch und massiv (um durchschnittlich knapp 16 Prozentpunkte) unterschätzt wurden. Fast zwei Drittel der Stichprobe machten keinen klaren Unterschied zwischen dem validesten und zweitvalidesten Cue. Methodologisch kann an der Arbeit von Hausmann et al. kritisiert werden, dass die Lernphase jeweils von den Einschätzungen der Validitäten unterbrochen wurde. Dieser Umstand lässt offen, wie die Autoren selbst bemerkten, ob die Einschätzung der Cue-Validität eher den momentanen Eindruck der Validität aus den letzten Durchgängen widerspiegelt oder das über das gesamte Experiment erworbene Wissen repräsentiert.

Im Grossen und Ganzen zeigen die berichteten Untersuchungen, dass die experimentelle Lernbarkeit der Cue-Güte Validität nicht sehr einfach, aber dennoch möglich ist. Diese Befunde legen argumentative Bausteine zur Rechtfertigung der Vorgehensweise, die Validitäten nicht nur mit Rangreihen, sondern gleich numerisch auf dem Information Board vorzugeben. Zum Beispiel gab Rieskamp (wie auch schon früher Rieskamp & Hoffrage, 1999) die Validitäten in dem Experiment „Ölsuche“ (2006, S. 1359) mit einer prozentualen „success probability“ und im Experiment „Kreditwürdigkeit“ mit „probability of success“ (Rieskamp & Otto, 2006, S.212), also einer Trefferwahrscheinlichkeit, vor. Rieskamp behielt diese Vorgehensweise der Validitätsvorgabe auch bei dem Experiment bei, in welchen die Versuchspersonen die Universität mit den meisten Publikationen finden mussten (Rieskamp & Hoffrage, 2008). In den neusten kritischen Publikationen zum PMM (Dougherty, Rick, & Franco-Watkins, 2008) wird jedoch festgestellt, dass es nicht weiter erstaunlich ist, dass die Informationssuche

meistens gemäss der Rangreihe der Validität verläuft, wenn diese Rangreihen von den Experimentatoren vorgegeben werden.

Die Diskriminationsrate von Cues

Die bis hierhin dargestellten Experimente beschränkten sich auf das Gütemass der Cue-Validitäten. Nun postulierten aber Martignon und Hoffrage schon 1999, dass ein Cue vor allem dann nützlich ist, wenn er – in einem Zwei-Alternativen-Fall – zwischen den beiden Optionen überhaupt diskriminiert. Diese Kombination der beiden Gütemasse Validität und Diskriminationsrate wurde als Success formuliert. Die Success-Strategie impliziert eine Informationssuche gemäss „Validität multipliziert mit der Diskriminationsrate“ unter der Voraussetzung, dass nicht weitergesucht wird, wenn dieser Cue nicht diskriminiert, sondern geraten werden muss. Die Suchstrategie Usefulness ist ebenfalls als „Validität multipliziert mit der Diskriminationsrate ($V \cdot D$)“ definiert. Aber im Gegensatz zu Success wird nicht von einem Stoppen der Informationssuche nach dem ersten Cue ausgegangen. Newell, Rakow, Weston und Shanks (2003) weiteten das Lernen sodann auch auf die Diskriminationsrate aus und liessen ihre Versuchspersonen die Hierarchie der Validitäts- und Diskriminationsrate im bekannten Börsenszenario in 64 Durchgängen lernen. Die Verhaltensdaten und die Fragebogendaten wurden von den Autoren so gedeutet, dass die Versuchspersonen fähig waren, die Cue-Gütemasse zu lernen, mit der Einschränkung, dass bei den numerischen Einschätzungen die Validitätsrate unterschätzt (von .09 bis .27) und die Diskriminationsrate überschätzt (.04 bis .16) wurden. Das Fazit dieser Forschungsergebnisse war dann: Können die Personen die Reihenfolge der Cues frei wählen, dann setzt sich die Suchregel Success durch (siehe auch Lagnado, Newell, Kahan, & Shanks, 2006; Rakow, Newell, Fayers, & Hersby, 2005). Dabei kritisierten Christen, Hausmann und Läge (2006), dass bei den dargestellten Ergebnissen die Hypothese von nicht gelernten Cue-Gütemasse nicht verworfen werden kann. Zwar entsprechen in der Untersuchung die subjektiv eingeschätzten Rangreihen den objektiven, aber der zur Testung verwendete Page-Test ist kein sehr sensibler Test. Zudem liessen die Autoren den Befund unerklärt, dass ca. 40% der Versuchspersonen meistens einen Cue wählten, der keiner der Suchstrategien von Validität, Diskriminationsrate oder Success entsprach. Christen et al. entwickelten daraufhin ein Lernexperiment, das sogenannte Äffchen-Experiment, in dessen Setting die Versuchspersonen die Validität und Diskriminationsrate nicht in einem Zwei-Alternativen-Fall, sondern Ein-Alternativen-Fall lernen mussten. Einerseits war die Lernphase eindeutig länger (100-200 Durchgänge) und andererseits war das Setting möglichst realitätsnah aufgebaut. Die Resultate in der Testphase zeigten, dass 21 (von 24 Versuchspersonen) die Diskriminationsrate und 18 die Validitätsrate im Verhaltenstest richtig einsetzten und 14 Personen die beiden Masse adäquat zu Success integrieren konnten.

Im Gegensatz zu den zahlreichen Experimenten zum Lernen von Cue-Validitäten ist die experimentelle Datenlage zum Lernen der Diskriminationsrate eher dünn gesät. Es konnte zwar teilweise ein Lernen und korrektes Anwenden der Diskriminationsrate gezeigt werden (Christen et al., 2006), aber teilweise häufig mit grossen Inkonsistenzen (Newell, Rakow, Weston, & Shanks, 2004). Dennoch argumentieren Läge et al. (2005), basierend auf den Erkenntnissen von Bröder zu der Lernbarkeit der Cue-Validität, dass man nicht nur die Validität, sondern auch die Diskriminationsrate numerisch auf dem Information Board vorgeben kann. So liessen sie die Cue-Güten Validität und Diskriminationsrate einerseits in einem Ex-

periment zu fiktiven chinesischen Städten und andererseits in dem Experiment „Millionärsquiz“ nicht mehr lernen, sondern stellten diese mit numerischen Zahlen direkt auf dem Information Board dar. Mit diesem methodischen Setting kamen die Autoren in den Befunden zu den fiktiven chinesischen Städten zu folgendem Schluss: „Klar ist nur: An der Validität allein orientieren sich die Versuchspersonen nicht, und sie haben auch kein individuell durchgehend angewendetes numerisches Modell, welches die Suchreihenfolge durch eine Kombination aus Validität und Diskriminationsrate erklärbar machen würde“ (S.11). Beim Millionärsquiz dektieren die Autoren bei etwa der Hälfte der Personen eine Informationssuche gemäss Validität, bei der anderen Hälfte gemäss Success. Aber im Allgemeinen sind die intraindividuellen Anwendungen der Suchregel nicht sehr konsistent. Die Autoren erklären dies folgendermassen: „Vielleicht sind Menschen gar nicht dazu in der Lage (wie u.a. von Kahneman & Tversky gezeigt), die beiden präsentierten Werte V und D hinsichtlich der postulierten Gütemasse Success und Usefulness korrekt zu verrechnen“ (S.16).

Werden Informationssuchstrategien anhand der Vorgabe der Cue-Gütemasse klassifiziert, stellt sich die Frage, ob die Ergebnisse womöglich ein Resultat einer reinen Verrechnungsaufgabe darstellen und diese Aufgabe die Personen zu einem „statistischen“ Entscheidungsmodus, wie von Bröder postuliert, verleitet. Sowohl bei der Validität wie auch der Diskriminationsrate ist es anhand der bis anhin erhobenen Verhaltensdaten schwierig zu testen, ob die Personen diese numerischen Angaben der Cue-Güte konzeptuell richtig verstanden haben und anwenden können oder ob sie diese Angaben bloss miteinander verrechnen und daher auch immer wieder einmal andere Strategien verfolgen. Bei dem Gütemass der Cue-Stabilität eröffnet sich nun aber die Möglichkeit, diese Problemstellung mit einigen Vorannahmen näher zu beleuchten.

Die Stabilitätsrate von Cues

Mit einem Experiment zu dem Gütemass der Cue-Stabilität (siehe ausführliche Beschreibung in Zurbriggen, Christen, Hausmann, & Läge, 2008) führten wir ein neues Gütemass ein. Wir nahmen an, dass in einer sich über die Zeit hin verändernden Umwelt auch Cues unterschiedlich stabil in ihrer Vorhersagefähigkeit sein müssen. Um das Experiment besser kontrollieren und auswerten zu können, gaben wir die Stabilitätsrate und die Validität der Cues numerisch auf einem Information Board vor und nahmen eine Erlernbarkeit der Stabilität a priori an. In dem Experiment wurden die Cue-Stabilitäten gegenläufig zu den Cue-Validitäten von vier Cues konstruiert, damit unterschiedliche Suchstrategien identifiziert werden konnten. Die Befunde zeigten, dass einige Personen auf die Cue-Stabilität reagierten und den stabilen Cue in ihrer Informationssuche integrierten. Die gefundene vermehrte Kaufhäufigkeit des stabilen Cues sagt prinzipiell jedoch wenig über das Verständnis der Personen zum Mass der Stabilität aus. Ein stabiler Cue macht eine Aussage über die prospektive Entwicklung der Zielvariablen. Daher macht es Sinn, diesen Cue zu kaufen, um die Informationssuche mit Blick in die Zukunft auszurichten. Mit der Überprüfung von zwei gezielten Hypothesen lässt sich untersuchen, ob die Personen diesen sehr stabilen Cue mit dieser Intention kauften, oder ob sie einfach den zusätzlich hohen numerischen Wert der Stabilität in ihre Informationssuche „statistisch“ integrierten:

1) Wir nehmen an, dass die Cue-Stabilität nicht verstanden wird, wenn der stabilste Cue nach einer Konsultation gleich nochmals konsultiert wird. Zudem nehmen wir an, dass die Stabilität der Information verstanden wird, wenn die Personen den stabilen Cue konsultieren, ohne ihn in der unmittelbaren Runde zuvor konsultiert zu haben.

2) Die Cuekaufhäufigkeit vom stabilsten Cue sollte mit der gesamten Cuekaufhäufigkeit zusammenhängen. Wenn die Personen, welche den stabilen Cue häufiger kaufen, dies mit Wissen um dessen zukünftigen Nutzen tun, dann sollte dies zu einer Abnahme der Gesamtkaufhäufigkeit führen.

Methoden und Resultate

Ausgewertet wurden die Daten aus der Bedingung mit Stabilitätsangaben der Cues vom Experiment zum Börsenszenario, welches ausführlich in Zurbiggen et al. (2008) beschrieben ist. Die Stichprobe umfasste 60 Versuchspersonen.

Es zeigte sich in dem erwähnten Experiment, dass der stabile Cue eindeutig häufiger konsultiert wird, wenn diese Cue-Stabilitätsangabe vorhanden ist als wenn diese Werte nicht angegeben werden. Ein sonst wenig valider Cue wird mit einer sehr hohen Stabilitätsangabe plötzlich häufiger konsultiert als ohne. Insgesamt wurde in der Bedingung mit den Stabilitätsangaben der stabilste Cue dreimal häufiger ($M=10.8$, $SD=10.4$) angeschaut als in der Bedingung ohne Stabilitätsangaben ($M=3.6$, $SD=5.1$; $t(59)=-5.2$, $p<.001$). Die hohe Standardabweichung deutet jedoch darauf hin, dass nur einige Versuchspersonen dieses Verhalten zeigten: von den 60 Versuchspersonen konsultieren 37 Personen den sehr stabilen Cue in der Bedingung mit den Stabilitätsangaben mehr als viermal.

Wird die Stabilitätsrate von den Versuchspersonen auch richtig verstanden und angewendet oder erhöht die bloße Angabe der Stabilität die Attraktivität dieses Cues? Von den 622mal, in denen der stabilste Cue angeschaut wurde, ist er in mehr als der Hälfte der Fälle, d.h. 345mal ($M= 5.8$, $SD= 9.7$) zweimal hintereinander konsultiert worden. 34 Personen suchten einmal oder nie den Cue zweimal hintereinander, 8 Personen zwei bis viermal und 18 Personen taten dies mehr als viermal. Weniger häufig, und zwar in knapp 30% der Fälle (184mal, $M= 3.1$, $SD= 3.3$)¹, erfolgte die Konsultation des Cues, ohne ihn vorher und nachher angeschaut zu haben, wobei 25 Personen dies fast gar nie taten (null bis einmal), 17 Personen zwei bis viermal dieses Muster zeigten und 18 Personen sich mehr als viermal so verhielten. Nur zwei Personen zeigten beide Suchmuster mehr als viermal. Der Unterschied zwischen den beiden Suchmustern wird im t-Test für abhängige Stichproben knapp signifikant ($t(59)=1.7$, $p(\text{einseitig}) <.05$). Wenn die Personen den stabilsten Cue kaufen, dann konsultieren sie ihn häufiger gleich zweimal hintereinander als mit Zeitabständen dazwischen.

Die Kaufhäufigkeit des stabilsten Cues und die Abnahme der Häufigkeit innerhalb der Person (im Vergleich zum Szenario ohne Stabilitätsangabe) korreliert sehr schwach mit $r=-.17$ (nicht signifikant). Das bedeutet, dass diejenigen Personen, welche den stabilsten Cue konsultierten, nicht zwingend weniger Informationen kauften.

¹ Die Summe von 345 und 184 entspricht nicht der Gesamthäufigkeit von 622 Käufen, weil einige Käufe (93mal) am Anfang und Ende der Bedingung keinem der beiden Muster zugeteilt werden können.

Diskussion

Um aufwändige Lernexperimente abzukürzen, ging man in neueren Experimenten dazu über, die Gütemasse der Cues numerisch vorzugeben. Bei der Validität und der Diskriminationsrate ist dieses Vorgehen noch einigermaßen gerechtfertigt, weil die Lernbarkeit der beiden Gütemasse – wenn auch mit Schwierigkeiten – doch gezeigt werden konnte. Dennoch fehlt auch hier der direkte Vergleich von dem Suchverhalten bei gelernten und vorgegebenen Cue-Gütemassen. Es müsste dringend ein Experiment durchgeführt werden, bei dem die internen Repräsentationen, welche die Versuchspersonen von den Cue-Gütemassen Validität und Diskriminationsrate sich bilden, intraindividuell untersucht werden und zwar in Abhängigkeit davon, ob diese Gütemasse sequentiell gelernt oder auf dem Information Board angegeben werden.

Da wir von einer Erlernbarkeit der Validität und der Cue-Stabilität ausgingen, gaben wir die Stabilitätsrate und die Validität der Cues numerisch auf dem Information Board vor. Auf den ersten Blick sah es dann auch so aus, als nähme zumindest ein Teil der Versuchspersonen in einer Reihe von Durchgängen auch Informationen für zukünftige Handelstage klar mit in den Blick, indem sie den stabilen Cue in ihrer Informationssuche mitberücksichtigten. Die hier dargestellten Auswertungen zeigen aber, dass nur wenige Probanden die Stabilität als eine weitere, neben der Validität bedeutsame Qualität probabilistischer Hinweiscues verstanden haben und Information suchen, die auch in Zukunft noch Wert besitzen werden. Mehrheitlich scheint die Angabe der Stabilität den Versuchspersonen zu indizieren, dass dieser Cue irgendwie wichtig sei, so dass doch etwas mehr als die Hälfte der Personen diesen auch konsultierten. Von diesen Personen haben aber nur ein paar wenige auch verstanden, was der zeitliche Nutzen des Cues ist und konsultieren ihn in unterschiedlichen Zeitabständen. Die meisten konsultieren den Cue gleich zweimal nacheinander und nutzen dessen Stabilitätsinformation somit nicht. Zudem muss bemerkt werden, dass dies nur eine Subgruppe ist: die andere Hälfte der Personen beachtet das Stabilitätsmass kaum, sondern fokussiert die Informationssuche auf die Validität der Cues.

Zusammenfassend kann gesagt werden, dass die aktuellen Befunde zeigen, dass die Situation bei vorgegebener Güte künstlich bleibt und im Falle der Stabilität für die Versuchspersonen anscheinend unverständlich. Dieser Befund verstärkt die Kritik an vorgegebenen Gütemassen in der experimentellen Forschung zur Informationssuche und legt für weitere experimentelle Untersuchung ein Lernen von Validität, Diskriminations-, und vorallem der Stabilitätsrate nahe. Insbesondere dann, wenn es um die Identifikation von verschiedenen Suchstrategien vor einer Entscheidung geht.

Literatur

- Bergert, F. B., & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and take-the-best model of decision making. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 33, 107-129.
- Bröder, A. (2000a). Assessing the empirical validity of the "Take The Best"-heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1332-1346.
- Bröder, A. (2000b). "Take the best - Ignore the rest" Wann entscheiden Menschen begrenzt rational? Lengerich: Pabst.
- Bröder, A. (2003). Decision making with the "adaptive toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 611-625.
- Bröder, A. (2005). Entscheiden mit der "adaptiven Werkzeugkiste". Ein empirisches Forschungsprogramm. Lengerich: Pabst Science Publishers.
- Bröder, A., & Gaissmaier, W. (2007). Sequential processing of cues in memory-based multi-attribute decisions. *Psychonomic Bulletin and Review*, 14, 895-900.
- Bröder, A., & Newell, B. (2008). Challenging some common beliefs: Empirical work within the adaptive toolbox metaphor. *Judgement and Decision Making*, 3(3), 205-214.
- Bröder, A., & Schiffer, S. (2003). Take The Best versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General*, 132, 277-293.
- Bröder, A., & Schiffer, S. (2006). Stimulus Format and Working Memory in Fast and Frugal Strategy Selection. *Journal of Behavioral Decision Making*, 19, 361-380.
- Christen, S., Hausmann, D., & Läge, D. (2006). Ist ein Gefühl für Informationsgüte lernbar? Ein Lernexperiment zu Validität, Diskriminationsrate und Success von Informationsquellen. *AKZ-Forschungsbericht 37*, Zürich: Angewandte Kognitionspsychologie.
- Dougherty, M. R., Rick, T., & Franco-Watkins, A. (2008). Postscript: Vague Heuristics Revisited. *Psychological Review*, 115(1), 211-213.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 506-528.
- Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In P. M. T. G. Gigerenzer, & the ABC Research Group (Ed.), *Simple heuristics that make us smart* (pp. 3-34). New York: Oxford University Press.
- Hausmann, D., Christen, S., & Läge, D. (2005). Und sie lernen es doch! Das Einschätzen der Validität von probabilistischen Hinweis-Cues und der Einfluss der subjektiven Repräsentation auf das Urteilsverhalten. *AKZ-Forschungsbericht 8*, Zürich: Angewandte Kognitionspsychologie.
- Hausmann, D., & Läge, D. (2005). Wenn Limiten zum Strategiewechsel führen: Wechsel der Suchstrategie unter situativen Einflüssen. *AKZ-Forschungsbericht 6*, Zürich: Angewandte Kognitionspsychologie.
- Läge, D., Hausmann, D., Christen, S., & Daub, S. (2005). Was macht einen "guten Cue" aus? Strategien der Informationssuche beim heuristischen Entscheiden unter Unsicherheit. *AKZ-Forschungsbericht 5*, Zürich: Angewandte Kognitionspsychologie.
- Lagnado, D. A., Newell, B. R., Kahan, S., & Shanks, D. R. (2006). Insight and Strategy in Multiple-Cue Learning. *Journal of Experimental Psychology: General*, 135(2), 162-183.
- Martignon, L., & Hoffrage, U. (1999). Why does One-Reason Decision Making Work? A case Study in Ecological Rationality. In G. Gigerenzer, P. Todd & A. R. Group (Eds.), *Simple Heuristics that make us smart*. New York, Oxford: Oxford University Press.
- Newell, B., Rakow, T., Weston, N., & Shanks, D. (2004). Search strategies in decision making: The success of 'success'. *Journal of Behavioral Decision Making*, 17, 117-137.
- Newell, B., & Shanks, D. (2003). Take the best or look at the rest ? Factors influencing one-reason decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 53-65.
- Newell, B., Weston, N., & Shanks, D. (2003). Empirical tests of a fast and frugal heuristic : not everyone "takes-the-best". *Organizational Behavior and Human Decision Processes*, 91, 82-96.
- Newell, B. R., Rakow, T., Weston, N., & Shanks, D. R. (2003). Search strategies in decision making: The success of 'success'. *Journal of Behavioral Decision Making*, 17, 117-137.
- Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, 16, 366-387.
- Rakow, T., Newell, B., Fayers, K., & Hersby, M. (2005). Evaluating Three Criteria for Establishing Cue-Search Hierarchies in Inferential Judgment. *Journal of Experimental Psychology / Learning, Memory & Cognition*, 31(5), 1088-1104.
- Rieskamp, J. (2006). Perspectives of Probabilistic Inferences: Reinforcement Learning and an Adaptive Network Compared. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(6), 1355-1370.

- Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics and how can we tell? In G. Gigerenzer, P. Todd & A. R. Group (Eds.), *Simple Heuristics That Make Us Smart* (pp. 141-168). New York: Oxford University Press.
- Rieskamp, J., & Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica*, 127, 258-276.
- Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135, 207-236.
- Zurbriggen, S., Christen, S., Hausmann, D., & Läge, D. (2008). Entscheiden in zeitlich dynamischen Umwelten: Suchen und Stoppen mit unterschiedlich stabilen Cues. *AKZ-Forschungsbericht 44*, Zürich: Angewandte Kognitionspsychologie.
- Zurbriggen, S., Hausmann, D., Christen, S., & Läge, D. (2007). Die zeitliche Dynamik von Umwelten: Heuristisches Wahlverhalten am Beispiel eines Trinkwasser-Szenarios mit und ohne Zeitachse. *AKZ-Forschungsbericht 41*, Zürich: Angewandte Kognitionspsychologie.

DANKSAGUNG

Als erstes gebührt mein Dank Damian Läge, der sich Zeit für unsere gemeinsame Arbeit nahm und mich in meiner methodischen Unbedürftigkeit geduldig unterstützte. Ohne die Mitarbeit von Stephan Christen und Daniel Hausmann-Thürig wären die Experimente nie so exakt durchdacht und genau umgesetzt worden. Wolfgang Marx danke ich für den Einblick in die Lehre und der Projektgruppe Urteilen und Entscheiden für das Mitdenken und Diskutieren. Auch meinen Bürospöndlis Dieter Sträuli, Roland Streule, Samy Egli, René Oberholzer, Matthias Häne und Stefan Ryf dank ich für die Stunden zwischendurch und Ursula Jäger für die aufmunternden Worte. Bedanken möchte ich mich auch bei Andreas Imhof, Marco Bleiker, Laura Jessica Schneider, Ina Sonogo und Patrick Leemann für die Unterstützung bei der Datenerhebung und allen Versuchspersonen, die ihre Zeit für die Teilnahme an den Experimenten investierten.

Persönlich möchte ich mich bei meiner Familie bedanken: meine Eltern Heidi und Franz-Josef Zurbriggen-Blum unterstützen mich in allem was ich anpacke und mein Bruder Sebastian Zurbriggen holte mich zudem einige Male früh aus den Federn. Meinen Freunden und Freundinnen sei allen gedankt für die Unterstützung im Leben neben der Diss, insbesondere Julia Blum, Karin Aregger, Janine Fuchs, Esther Grosjean, Ruth und Edim Suljadzic, Mia Meyer, Leana Schmid, Marie-Christine Abgottspon, Petra Schneider und Dario Carrozza. Zudem danke ich auch der infoSekta und deren Team, die mir das Sammeln von praktischer Berufserfahrung ermöglichte.

CURRICULUM VITAE



Seraphina Zurbriggen
Giesshübelstr. 94
8045 Zürich
079 / 727 31 68
serazub@hotmail.com

Jahrgang 16.2.1976
Heimatort Betten / VS

Ausbildung Doktorandin an der Abteilung für Angewandte Kognitionspsychologie (2005-2008) an der Universität Zürich und Assistentin an der Abteilung für Allgemeine Psychologie/ Kognition an der Universität Zürich (2007-2008).

Studium im Hauptfach Psychologie, Nebenfächer Psychopathologie und Filmwissenschaften an der Universität Zürich (1998-2004).

Berufliche Erfahrungen Fachstelle für Sektenfragen, infoSakta: Mitarbeiterin (seit 2005).

Institut für Arbeitspsychologie (Ifap) der ETH Zürich: Wissenschaftliche Hilfskraft (2003).

Praktika Kantonale Psychiatrie Rheinau: Psychopathologie für Erwachsene (2001).

Austin State Hospital, USA: Psychopathologie bei gehörlosen und behinderten Menschen (2000).